CONTENTS

From the Editor’s Desk 3
The Catalog as Access Mechanism: Background and Concepts.
    Patrick Wilson 4
Technology and the Online Catalog. Peter S. Graham 18
Characteristics of Online Public Catalogs. Stephen R. Salmon 36
The Human Dimension of the Catalog: Concepts and
    Constraints in Information Seeking. Leigh Estabrook 68
Use of Classification in Online Retrieval. Elaine Svenonius 76
Practical Considerations of the Current Capabilities of Subject
    Access in Online Public Catalogs. Anne Grodzins Lipow 81
Futuristic Aspects of Subject Access. Phyllis A. Richmond 88
Letters to the Editor 105
For the Record (Annual Report of the Decimal Classification
Instructions to Authors 111
EDITORIAL BOARD

Editor, and Chairperson of the Editorial Board .................ELIZABETH L. TATE

Assistant Editors:

PHYLLIS A. RICHMOND ...............for Cataloging and Classification Section
EDWARD SWANSON

CAROLYN C. MOWROR ...............for Preservation of Library Materials Section
FRANCIS F. SPREITZER ...............for Reproduction of Library Materials Section
J. MICHAEL BRUER ...............for Resources Section
LINDA SAPP ..................for Serials Section

Editorial Adviser:

DORIS H. CLACK (for Regional Groups)

Liaison with RTSD Newsletter:

ARNOLD HIRSHON, RTSD Newsletter Editor

Library Resources & Technical Services (ISSN 0024-2527), the quarterly official publication of the Resources and Technical Services Division of the American Library Association, is published at 50 E. Huron St., Chicago, IL 60611. Business Office: American Library Association, 50 E. Huron St., Chicago, IL 60611. Advertising Traffic Coordinator: Leona Swiech, Central Production Unit, ALA Headquarters, 50 E. Huron St., Chicago, IL 60611. Circulation and Production: Central Production Unit/Journals, ALA Headquarters, 50 E. Huron St., Chicago, IL 60611. Subscription Price: to members of the ALA Resources and Technical Services Division, $10 per year, included in the membership dues; to nonmembers, $20 per year; single copies, $5.

Second-class postage paid at Chicago, Illinois, and at additional mailing offices. POSTMASTER: Send address changes to Library Resources & Technical Services, 50 E. Huron St., Chicago, IL 60611.

Library Resources & Technical Services is indexed in Library Literature, Library & Information Science Abstracts, Current Index to Journals in Education, Science Citation Index, and Hospital Literature Index. Contents are listed in CALL (Current Awareness—Library Literature). Its reviews are included in Book Review Digest, Book Review Index, and Review of Reviews.

Copies of books for review should be addressed to Arnold Hirshon, Editor, RTSD Newsletter, Duke University, Box 9184—Duke Station, Durham, NC 27706. Do not send journal issues or journal articles for review.

The contents of this journal, unless otherwise indicated, are copyrighted by the Association. All material in this journal copyrighted by the American Library Association may be photocopied for the noncommercial purpose of scientific or educational advancement.

© American Library Association 1983

Publication in Library Resources & Technical Services does not imply official endorsement by the Resources and Technical Services Division nor by ALA, and the assumption of editorial responsibility is not to be construed necessarily as endorsement of the opinions expressed by individual contributors.
From the Editor's Desk

Two entries from the twenty-five-year cumulated index to *LRTS* succinctly summarize the cataloging history of the past few decades:

Card catalogs:
- Closing: 25: 186-95
- Expansion: 1: 104-8

The following entry in a *LRTS* fifty-year cumulated index will tell the story of this issue and of the decades to come:

Online catalogs: 27:4-93

In this issue you will find four of the papers given at the recent preconference, "Prospects for the Online Catalog," and three papers from the conference program entitled "Subject Analysis in the Online Environment." All but one of these papers appear essentially as they were presented in Philadelphia, with only minor editorial changes to eliminate remarks of merely transitory relevance, provide documentation, and achieve consistency in matters of capitalization, spelling, etc. The paper by Stephen R. Salmon, however, has had repeated revisions to incorporate the responses to his questionnaire that arrived after the preconference; therefore, "Characteristics of Online Public Catalogs" is as up to date as author and editor can make it. The article by Hines and Winkel nicely complements the conference and preconference papers, for it, too, examines the question of subject access and proposes a computer-oriented solution.

In the next issue we plan to publish a selection of papers from the conference-within-a-conference, "Research by and for Librarians: The Needs, the Methods, the Opportunities," along with some reports of original research by and for technical service librarians.—*Elizabeth L. Tate.*
I want to consider what catalogs are supposed to do; how and to what extent they do it; and what differences it makes to our notions of goals and methods when we think of an online catalog instead of a card, book, or microform catalog. I take the standpoint not of a cataloger, which I'm not, but of one interested in the whole bibliographical universe and the various ways in which we try to organize and exploit it. By bibliographical universe, I mean the whole body of public records of talking and writing. Published records are central, but we can consider unpublished records as part of the bibliographical universe if they are accessible to the public. Records of activities other than talking and writing—of painting and drawing and composing music, for example—are outside my scope. Insofar as people's talking and writing has permanent effect, it does so in two quite different ways, one very concrete and one quite abstract. Concretely, some marks are made on paper, or some traces left on a piece of magnetic tape; abstractly, a certain string of words has been produced—a certain text. For historical purposes the original manuscript or other original physical record is of enormous importance, but from another standpoint, the crucial thing about what one says or writes down is simply that it specifies a definite sequence of words, a definite text, that can appear again and again in different physical manifestations. The words are not likely to be new; only the particular choice and order are new—and they may not be very new. The text may incorporate strings of words that resulted from one's own or other people's prior performances; it may be almost nothing but others' strings of words, cut up and stuck end to end with a bit of transition supplied. The first published appearance of the text, the first edition, if there is one, may be the only appearance; for most texts the first edition is also the last. If this were so

Patrick Wilson, professor, School of Library and Information Studies, University of California at Berkeley, presented this paper as the opening address at the "Prospects for the Online Catalog" preconference, held in Philadelphia on July 8-9, 1982. The author wishes to acknowledge the helpful comments on early drafts from Julia Cooke and Theodora Hodges and to express his appreciation for the valuable advice he received from Raynard C. Swank.
for all texts, life would be easier for catalogers. But the first appearance may be just the beginning. The same text may appear again and again, separately or as part of a longer string, as in an anthology. And the text may serve as the basis for the production of new related texts that constitute new *versions* of the old text: rethinking and correcting mistakes lead to revised versions of the initial text; expansion and elaboration to enlarged versions; simplification and contraction to abridged versions; recasting in other styles or literary genres to adaptations; addition by someone else of critical apparatus to critical editions; and, of course, translation into other languages. Others may produce their own texts that expound the first text, that imitate it, that summarize and criticize it, that attempt to refute or defend or improve on it, and so on indefinitely. These can be called *derivatives* from the first text. Once in the bibliographical universe, a text can become an ancestor of a huge family of other texts related to it in an immense variety of ways and degrees; and any of these near and remote relatives, versions and derivatives, might have one or many published appearances, and start its own family of related texts. This universe of texts, each making one or more published appearances and in one or more versions, is the world to which the catalog is to give entry.

You’ll recognize this talk of text and published appearance as something like the familiar distinction between *literary unit* and *bibliographical unit*, or between a *book* and the *work* it contains. Some such distinction is inescapable. I prefer to talk about texts and related texts: versions and derivatives, than to talk of works or literary units. I know how to tell if two publications contain the same text or not—they do if they contain the same sequence of words—but I’m not sure I know how to tell if two publications contain the same work or not. The trouble is that we seem to be required to make an either/or decision: either two texts are texts of the same work or they are not. But the relations among texts are endlessly complex; there are so many possible gradations and shadings of relationship, that I find it quite hopeless to look for a general rule or principle according to which I could make that either/or decision. How much can a text be revised before it becomes a different work? How exact does a translation have to be to count as a translation of the same work? I think there is no general way of answering such questions, and thus doubt the possibility of clarifying the notion of a work. There may be a useful distinction to be made between work and text, but I don’t know what it is. For now, we’ll simply do without the notion of a work.

The bibliographical universe is in a certain sense a self-organizing world: there are organizing devices internal to the world. We think most readily of the big independent periodical indexes, abstracting services, published catalogs, national bibliographies, comprehensive subject bibliographies as constituting the bibliographical apparatus. These are designed to be instruments of bibliographical control, by which I mean instruments of discovery, allowing first the discovery of the very existence of a text or book and then the discovery of where copies can be found. The bibliographical universe is not under effective bibliographical control until anyone can discover those of its inhabitants that will suit his or
her purposes, whatever those purposes may be. A serious explorer of the bibliographical universe may manage to avoid use of those big independent bibliographical works, relying instead on the network of references within the nonbibliographical literature; there is more than one way to skin a cat, and the big formal bibliographical works are not the only means of discovery. One might have satisfactory bibliographical control of the universe of texts without relying on the formal bibliographical apparatus at all. But in fact that apparatus is often essential as a means of discovery of other elements in that universe. Now just what place is the catalog supposed to have in the whole array of bibliographical instruments? The catalog of a particular library is just one more piece of bibliography among thousands of others. What is it supposed to do that the other bibliographical works don’t, or don’t do so well? What does it do that wouldn’t be done, or wouldn’t be done so well, if you didn’t have it but did have everything else?

If we want to know what a library catalog is supposed to do, the best place to start is Cutter’s classic statement of the “objects” of a catalog, which has been repeated with minor modifications by practically everyone who discusses the nature of the catalog. Cutter says, and everybody agrees, that the first objective is to enable a person to find a book of which either the author or the title or the subject is known. Three questions arise in my mind at once. First: seriously, is the catalog to enable us to find a book? Then it should tell us where the book is. Up to now, it has not generally done so. It has given a theoretical location: a call number. But the books that I want seem rarely to be at their theoretical locations; they’re on loan, in process, lost, in storage, misshelved, or waiting to be reshelved. To find the actual location, one has to consult circulation records which, with luck, will tell one where the book actually is or, at least, when it is expected to be returned to the library. Up to now, one has often had to consult at least two files to find out where a book was; the catalog gave only the theoretical location. The situation changes when the catalog goes online; now the two separate files could be merged, and the single catalog could tell you, if not precisely where in space the book was, at least what its current status was. Now the catalog would come closer to being able to do what it was supposed to do.

But which books are these that the catalog is supposed to enable me to find? The books in the library? No, of course not, the books that the library has. But is that quite right? Isn’t it somewhat narrow-minded? Isn’t it true that what happens to be “owned” by a particular institution is only a part of what’s actually available at that place? The library can supply me with a book it doesn’t own but borrows from another library. If the library owns a book but it’s not actually available and won’t be available for weeks or months, it may be able to provide me with a copy that it doesn’t own. I recently read an LC copy of a book, all the local copies of which had disappeared. Why limit the catalog to things owned but perhaps not actually available? Why not extend it to cover things available through the library, whether or not owned by the library? The fact that things available from other sources are not instantly available is nothing to the point, for half the books the library owns that I want
aren’t instantly available either. Only the noncirculating library can come close to guaranteeing that the things the library owns will be instantly available in their theoretical locations; but even it can make available things it doesn’t own, after some delay. Why tie the catalog to the notion of ownership? One good reason is that almost everything in the bibliographical universe is potentially available at any library; if the interlibrary lending system is working as it should, I ought to be able to get almost anything after perhaps a long wait. And so the local catalog should list everything that anyone is willing to lend; and that would make it a big catalog, impractically big. But does that mean an impossible catalog in principle? Surely not; a stratified catalog, divided into zones of accessibility, in which you first look in the local zone, then move to a next-proximate zone, then to a further one, is in principle perfectly possible, and would be a better tool for finding a book than present catalogs limited to what’s locally owned.

Finally: Find a book? Why a book? Well, of course we want to enable people to find phonograph records and newspapers and so on; the catalog need not be limited to books. But that’s not the point. Books are not the only things identifiable by author and title and subject; so are the texts that books contain. I may know perfectly well what text I want without knowing what book or books, if any, it appears in. Don’t we want to say that the catalog is meant to allow one to find copies of texts? You can’t just find a text; it has to appear in some physical form. But if it’s the text that one is looking for, and if the same text appears in different books or in nonbook formats, one may not care at all which published appearance one gets so long as it contains the right text. This is a real difference, not just a question of words; a good book-finder may not be a good text-finder.

While it sounds good to say the catalog is supposed to enable us to find copies of texts, in fact our catalogs are not designed to do so except in a limited variety of cases. Very roughly, it will do that for long texts but not for short texts. If the text I want happens to occupy an entire book, I can find it in the catalog, provided I can recognize from the description that it is the text I’m looking for. If it occupies only part of a book, I may be able to find it in the catalog; I can do so if this part has been separately listed in the catalog (that is, if an analytic entry has been made), or if I find a description of the book that enumerates its contents. But unless the book is a collection of texts by the same author, I may never find it in the first place. Contents notes don’t help if one can’t find the catalog record in the first place. We don’t set ourselves the goal of allowing discovery in the catalog of periodical articles, single contributions in collections, single speeches in Congress, or the like. That is what periodical indexes and other bibliographical works are for. The basic unit for the catalog is not the text but the separate publication, the bibliographically independent publication. Our basic unit is indeed the book, when it’s not the serial publication treated as a single multivolume book. As far as texts are concerned, we’re mostly concerned only with texts that happen to occupy all or most of a book, and since books can be small or large, a short text may after all appear in the catalog and a longer text not ap-
pear. So if we happen to have two published appearances of the same text, one as the sole content of a book, and one as part of an issue of a periodical, or in several issues of a periodical, the first appearance will show up in the catalog but the second one won’t. Our choice of basic cataloging unit has the consequence that enormous numbers of texts are not locatable through the catalog without first using some other bibliographical work, to find what book or serial they are in. And it has the consequence that of the different copies of the same text in the library, some will be discoverable directly through the catalog, others discoverable only by an indirect approach through other bibliographical works. This is the standard practice, and I don’t quarrel with it; but I do want to point out that it makes a cleavage of the bibliographical universe that is hard to describe briefly and accurately, and that must be quite mysterious and even irrational to the unindoctrinated user of the library.

The strangeness of the situation becomes more striking when we turn to another group of objectives of the catalog. In Cutter’s formulation, the catalog is to show what the library has by a given author, or on a given subject, or in a given kind of literature. Note the exact wording: the catalog is to show what the library has. This is appealingly direct and to the point. It’s exactly what one might have expected a catalog to do. Cutter didn’t say that the catalog was to show what books a library had, but simply: what it had. He didn’t say the catalog was to help one find what the library had, but to show what it had. This objective has been quoted approvingly for a hundred years, but we don’t really mean it. It’s not our real objective—not, at least, if the objective is understood as it normally would be understood, to mean that one meant to show all that one had. We certainly don’t reveal all that we have by a particular author, and don’t even aim to do so. We don’t show all that we have on particular subjects, and don’t even aim to do so. In the case of subjects, it’s not just a matter of ignoring relatively short texts. Showing what you have means being explicit; if part of a book discusses a certain subject, and if you describe the book using terms applicable only to the whole book, you’re not being explicit about the part. If the part discusses, say, the topic of metaphor, but the book is described as a whole as being about language, you haven’t shown that you have that discussion of metaphor. We don’t aim to show, to reveal explicitly, all the separate or distinguishable contents of a book.

This is all right, or at least understandable. But it has odd consequences. A naive user might legitimately expect that, if you’ve got two copies of the same text in your library, you’ll list both; you certainly wouldn’t list one and hide the other. This is such a basic and intuitively reasonable expectation that we might call it a principle: the Principle of Parity, say. The same text appearing in different circumstances of publication should be treated the same in all of its appearances, at least to this extent: if you show one, show them all. But we don’t observe this principle. Cutter himself did have a rule requiring analytic entry for any part of a book that also appeared as a separate; we’ve forgotten that rule.

Cutter’s rules themselves appeared as an appendix in a government report on public libraries; they’ve also appeared separately, in several ver-
sions, as the entire content of a small book. The former appearance will be ignored by the catalog, the latter will be recognized. Parity is not a cataloger’s principle.

The same Principle of Parity applies to subject description, though it can take any of several forms. For example: if you have two texts that give the same amount of information about a subject, then show both if you show either; don’t show one and hide the other. Alternatively: if you have two texts of equal importance for the study of a subject, show both if you show either; don’t show one and conceal the fact that you have another equally important one. That has a corollary: for heaven’s sake, don’t list a small inferior work and conceal a large important work. Now you may immediately object: that’s a matter of value judgments, and we can’t expect the catalog to reflect value judgments (except insofar as it reflects value judgments involved in book-selection decisions). If anything has got to be value free and neutral, it’s the catalog. Just so: and for just that reason, the catalog is unavoidably a terrible guide to what the library has on a subject. It grossly violates the Principle of Parity in this evaluative version, but nothing can be done about that because to do anything would require evaluation, and we don’t evaluate. But the other version of the Principle, in terms of quantity of information (or quantity of text) is also grossly violated: a slight monograph will be listed, a huge treatise that incidentally has twice the amount of material on the same topic will go unrecognized as having anything to say on the topic. This is a consequence of the basic rule of subject cataloging, the rule of specific entry: a work all about a large subject is so described, the smaller subjects included in its scope are not explicitly shown. The rule is a sensible one, but adherence to it automatically results in violations of the Principle of Parity. We can of course say that we expect any intelligent person to realize that not all of the material on a subject will be in books of which that is the subject, and that he or she will have to look elsewhere to find what may actually be better for the purpose at hand. And to help in that further search, we provide a systematic set of cross-references that can be followed up. This is so, and a persistent and ingenious catalog user can make good use of that apparatus of cross-references. But providing some help to the persistent and ingenious catalog user is a far cry from showing explicitly what the library has. The catalog is selective, first in what it exhibits at all, second in what it tells about the things it does exhibit. The selectivity is defensible, but we really must not keep saying simply that the catalog shows what we have by an author or on a topic: it shows some of what we have. We’ll come back later to the question of whether it ought to show more.

The big questions to answer when designing a bibliographical instrument are these: first, what’s to count as one item, and what items, so counted, are to be included (the question of unit and scope); second, what information is to be supplied about each item; third, what information will allow one to find an item—how things are to be accessible, how organized, indexed, arranged (the question of access). We’ve been discussing the first of these, unit and scope; let’s turn to the others.

The big difference between conventional card or book or microform
catalogs, on the one hand, and online catalogs, on the other, is in access and display of information. Until now, decisions about access took the form of decisions about entries: main entry, added entries. An item appeared at a given number of places in the catalog—under the name of the author, under the title, under the name of a series, for example. At each appearance, the same descriptive information might be given, or, as was the case when Cutter was writing, the full information might be given only once, and only partial information given otherwise. That tedious old notion of the main entry made sense then; the main entry was the entry giving full information. With the advent of the unit card system, the notion dwindled to a ghost of its former self, as the notion of the entry you would make if you were going to make only one. (This is an oversimplification, but not by much.) But in a computer environment the notion of main entry seems to lose all meaning. And the notion of entry itself is transformed. The question is not where and how an item is to appear in a fixed list, but how it can be made to appear: first, what signals will produce a record, and second, what different forms the record can be made to take when it appears.

To emphasize how the notion of entry changes in a computer environment, let me remind you of the OCLC search keys. (In fact, it seems to me that the independent development by OCLC of its search keys was decisive in altering the notion of the entry.) The sequence of letters “bell, dean” is one way of making records for Saul Bellow’s novel The Dean’s December appear; another way is “dea,de,,”. It sounds very odd to say that Bellow’s novel is entered in that OCLC catalog under “bell, dean” and under “dea,de,,” but that’s in fact what it amounts to. Search keys or search statements or, as we might say, cues that will retrieve a record are the computerized catalog’s analogs or equivalents of entries in manual catalogs. The record for an item may be physically stored in just one place, but that one place is not the item’s one entry; it’s not an entry at all but the source of entries. Or rather: the question that used to be put as one of selection of main and added entries becomes transformed into the question of permissible search requests, of the kinds of instructions to a machine that will make a particular item appear. Simultaneously, the question of how much information to provide for each entry changes; the possibility of asking for different amounts of information (short, medium, full, as Cutter would say) means that the question is now what different amounts and formats of information we will allow the user to make appear. That finally does in the notion of main entry, for if one has the option of getting full information no matter what search request or cue one uses to make the item appear, then every entry is main entry, and the notion of main entry loses all point.

Instead of deciding what entries to make, then, we have to decide what kinds of search keys or search statements or cues and combinations of cues will suffice to retrieve a record. The great practical difference that the computer makes is (1) in the enormous increase it allows in the number of different elements of information that might be used as cues; (2) in the possibility of combining cues using the Boolean operators and, or, and not; and (3) in the possibility of automatically changing the access
system without altering the basic bibliographical records themselves. There are so many new possibilities that if a system designer were to ask: Which of the possible kinds of search capabilities would you like us to build into the system?, it’s hard to imagine how one could resist saying: We want them all, now. Naturally, we’ll want to be able to do what we could do in the old catalogs: if there was a point in making entries under author’s name and title and editor and series in the old catalogs, the point is presumably still valid, and we’ll want to be able to make a record appear by specifying author or title or editor or series. It would be odd and disappointing to find that an online catalog gave us less access than did the old system of main and added entries. But we want to be able to retrieve a record using partial information and combinations of partial cues: author’s surname only (instead of full name) plus some word or words from the title (not the exact title with all words in the right order); those cues plus publication data, or plus approximate publication date; part of a surname plus other cues; and so on. We want to be able to exclude things: to ask for works not written in Bulgarian, not published after 1970, and so on. The question is, How much of the full combinatorial capacity of the computer, working on what part of the range of elements in the full description, are we to make use of? If that is a radical transformation of our old question about how many entries to make for a particular record, it is still the same question at heart: how to make the records accessible. I can think of no general principle that one could use to help decide that question except the Principle of Generosity: the more the merrier. Or, to put it more solemnly: maximize the chances of finding what one is looking for, given the information one starts with. Minimize the restrictions on use of whatever information one has. If we take that line, the question of access converges on the question of description; if you’re going to supply a piece of information in the full description of an item, why not let one who happens to have that piece of information use it in trying to get that item to appear? If it’s worth giving at all, why isn’t it worth using for retrieval purposes, as one part of a search request? Clearly not very many of the possible kinds of complex search requests are going to be used often; simple search requests using authors’ names and partial titles are going to be the popular ones, I would expect. But from the point of view of the user, not the system designer, the desirable thing would be to allow the user to use whatever information he or she had to start with, in trying to locate particular items: to maximize accessibility.

I want to interrupt the story at this point to caution against too much satisfaction about the possibility of altering the access system automatically without changing the basic records. The passage of time causes problems for a catalog, some of which are all too familiar, some of which are, I think, unduly neglected. People change their names; organizations change their names too, as do serial publications. New versions of old texts appear. New things are written about. Old things get new names, terminology changes. This last kind of change is one that looks to be easy to cope with in an online environment; with automated authority control, changes of terminology are easily dealt with. But there is a
kind of change that cannot be accommodated so easily. The establish-
ment of a new subject takes time; subject catalogers wisely wait for a
while to let terminology settle down, though the easier it is to change
terminology, the less reason there is for waiting. But during the wait,
and even before one recognizes that there is something to wait for, peo-
ple have been producing books that in retrospect can be seen to be best
described by terminology that was not then available. After you finally
recognize the category of, say, popular culture as something that can be
the subject of a book, you might ask: How have we been describing
books on this subject up to now, when we lacked that subject heading?
The introduction of a new category, not simply a new name for an old
category, should be the occasion for a retrospective look at what has pre-
viously been done, for the chances are that some of the old work is in
error, given the new vocabulary. Recataloging is in order. But, so far as
I know, it is not established practice to reexamine old cataloging in the
light of new vocabulary. Corrections may be made for accidental rea-
sions: a new edition of an old work, a need to correct old author entries.
But systematic review of old subject cataloging, I think not. Now as ter-
minology changes and new categories are introduced, over time the
amount of misdescribed material is bound to increase; the accuracy of
the subject catalog declines, the quality is gradually degraded. This is
something that automatic procedures cannot eliminate. There is no au-
tomatic recognition process for misdescribed works.

We’ve got rid of the notion of entry or, rather, replaced it with the
notion of permissible search request. In particular, we’ve got rid of main
entry, and so are rid of those wearisome debates about whether main
entry should always be under title, and those difficult decisions in partic-
ular cases about whether main entry should be under personal or corpo-
rate author. Good riddance! Now, might not the notion of authorship be
thrown out too? Mr. Gorman has recently managed somehow to per-
suade the library profession that corporate authorship is a bad idea and
should be entirely, or almost entirely, discarded (except for certain ‘em-
anations’ from corporate bodies). That was a most surprising feat. It oc-
curs to me that a profession that can do that ought to be able to take the
next small step and toss out the idea of personal authorship as well. I
hasten to say that I’m not suggesting abandonment of authority control.
We want to be able to pose search requests using personal names with
the assurance that we will retrieve things appearing with any form of the
name of the same person, and that requires standardization on names,
or links among different forms of name used by the same person, that
might be automatically activated when any particular form is used in a
search request. But we might keep authority control while abandoning
authorship, by which I mean, ceasing to do anything to indicate explic-
itly that a person is the author of a particular text. This will, I hope,
sound shocking; are not our catalogs firmly based on the authorship
principle? If we want to allow a person to find a book of known author-
ship, how can we do so without identifying authors as authors? In fact,
however, I think our catalogs are not firmly based on the authorship
principle. We do not really inquire into the question of whether a per-
sonal name on the title page is the name of the author of the book. We take things at face value. Authorship really means purported authorship. When we learn that a text purportedly written by a late president of the United States was actually written by his assistants, we don’t rush out to change the catalog records; that president is still the purported author. The managed textbooks that are written by anonymous hacks and that appear under the names of eminent professors would be listed, with a clear conscience, under their title-page purported authors. We recognize pro forma authorship, as when the twentieth edition of a text containing hardly a word set down by the original author is still ascribed to the original author, depending on the appearance of the title page. We’re not really serious about authorship, and that is quite right; title-page names are the names likely to be used by those looking for the texts, and hence are the ones to worry about. No, we are not serious about authorship; and why take seriously the job of establishing purported authorship? A bibliographical record might now simply contain a title-page transcription, and the proper names contained in that transcription made searchable; this might easily be done without taking any position on the question of whether a name is the name of the author. What difference would it make? We would allow people to search by specifying a personal name alone, or personal name and title; in either case those things (and maybe more) would be retrieved that would have been retrieved if we allowed them to search for items of which some person was the author. We can, I say to forestall instant objection, distinguish between personal names as subjects and personal names as nonsubjects, “bibliographical” personal names, all without explicitly indicating authorship. As far as I can see, dropping explicit indication of authorship would be no great loss. But it may be helpful and thus desirable to specify different functions or roles played by individuals, and allowing one to search for works of which an individual is the purported author without also retrieving those of which he is the purported editor or translator or introduction writer. The Principle of Generosity would favor this. Still, it is worth reflecting on how little we do in fact depend on the notion of authorship. We proceed, more and more, on the basis of quite another principle: the Principle of Maximum Feasible Superficiality. As much as possible, we take things at face value. I do not say this is a defect.

Let us return for a moment to corporate authorship. Absent the need to decide on main entry, the debate over corporate main entry disappears. If we drop the notion of authorship, the notion of corporate authorship disappears as well. As with personal names, corporate names appearing in a title transcription or elsewhere might be made searchable without explicitly indicating any particular functional relationship between corporate body and text. But if we want to assign function or role indicators to personal names, distinguishing purported authors from purported editors and the like, why should we not assign them to corporate names too? In particular, the role indicator of purported authorship? For this good reason, it will be said, that an organization as an entity cannot write a book; only people, singly or in groups, can write books. I suspect that the rejection of corporate authorship is in part
based on a romantic view of personal authorship. In scientific journals, for instance, the appearance of a personal name at the head of a journal article does not imply, hardly begins to suggest, that the person named did any writing; it only implies that the person had some unspecified role in doing the piece of research that is reported in the article. Mary Leakey wanted her name taken off a paper not because she hadn’t actually written any of the work (of course she hadn’t; she had lent the actual writers some bones), but because she disapproved of the content. Authorship no longer really implies writing, and purported authorship needn’t imply purported writing. The ways in which an organization might very seriously be eligible for authorship are explored deeply in Michael Carpenter’s book *Corporate Authorship;* ¹ I want to supplement that book by arguing, briefly, that since we don’t take personal authorship seriously (and shouldn’t), we needn’t take corporate authorship seriously either, but we needn’t take it any less seriously. For us, in practice, authorship is mainly a matter of name prominently displayed on title page; and it needn’t make any difference whether the name looks like the name of an individual, or a committee, or a project, or an event, or a corporation, or a government department. To attribute a role to an organization or group need no more imply any theory about what part different members of the group played in the putting together of a string of words, than attribution of the role of author to an individual has to imply a theory about what the person actually did. In other words: role attribution without commitment—superficial role attribution. The notion of corporate authorship is not dead; if one wants a deep analysis of what it seriously means, that’s available; without deep analysis, we are justified in using the notion as superficially as we use the notion of personal authorship.

Let us return to the question of subject description. One of the first things one is likely to think of, when contemplating a computer-based catalog, is that now we should take the occasion to begin revealing more of the content of our collections. Subject analysis is shallow; it often amounts to little more than paraphrasing the titles of books in a formalized vocabulary and syntax. The vocabulary and syntax are that of the Library of Congress subject heading system; they needn’t, of course, be that. There are plenty of alternative ways of formulating short formalized descriptions of content, the PRECIS system being the most talked about. But the point is that whatever system you employ, short descriptions of overall subject matter just don’t reveal much explicitly about content. And so one might say: now is the time to deepen the indexing, to make more of the content of the texts directly accessible. Lots of alternatives suggest themselves; for instance, a layered content description: first an overall summary of content; then a set of descriptions of separate parts or aspects of the content; then a further set of descriptions of smaller parts or aspects. Given a search procedure that allowed one to search in different layers, sticking to the top layer if one wanted, going deeper if one wanted, this would greatly increase retrieval capability, it would enlarge the amount of material on a topic that was explicitly shown as such. Should we not be working on plans to do just this?
I wouldn’t try to prevent anyone from producing a more deeply indexed subject catalog, but I’d like to argue against doing it. My reasons have to do, first, with the kind of bibliographical control that is most important to library users, and second, with the way we ought to be thinking about the subject component of the catalog. Let’s look at the user first. For most users most of the time, I suspect that it is fair to say that, when they’re not looking for a particular item they already know they want to see, they’re looking for something good. Years ago, I put it this way: that the kind of power one would most like to have is power to get the best textual means to one’s particular goal or end. Years of trying to find reasons to doubt that have left me still of that belief. Light users of libraries want something they will find good to read; serious users want the ones that will serve them best to accomplish some purpose. I wanted recently to catch up on what had been going on in the study known as behavioral decision theory; of course what I wanted was a good survey of recent work: authoritative, up to date, clear, well written, by someone who was fair to all parties, and so on. I wanted the best account I could get. The kind of bibliographical control I wanted was not the ability to discover everything there is that fits the bare topical description of what I’m looking for, but the ability to get the good ones and be spared the bad ones. There’s another and related matter: as often as not, what I’m looking for is something I can describe only in functional terms—I want something that would help me see or understand or solve something. I’d like to be able to say what’s bothering me, and then ask: What is there that would help me in this situation? I don’t care what the subject matter is, who the author is, how old it is, I will take anything that will be of help. I don’t know where to look in a list of subjects, for as far as I know the things that might help might be anywhere.

If I’m right about this, then the two things that make a text most interesting to me are things that we won’t or can’t provide in the catalog. We won’t say anything about value; even if we did, we might well say the wrong thing, for the value in question is relative to an individual’s situation and purposes. The other point is harder to grasp, I think. From the point of view of the user who wants materials that will help solve or at least clarify his problem, the cataloger approaches materials from the wrong end. The catalogers go book by book, and about each one ask: What is this particular book about? But to be of real help, they’d have to start not with particular books but with particular questions or problems, and ask about each book, What if anything might this book contribute to solving or clarifying this particular problem? Doing that would be an effort to organize materials functionally, by their relationship to particular problems. The cataloger doesn’t and can’t do that, has rather to take each book in sequence and describe it internally, and consequently produces an instrument that is fatally flawed from the serious user’s point of view.

We can’t provide evaluations, and can’t organize materials functionally, in terms of uses to which they can be put rather than topics they’re about. We’re stuck with content description in terms of subject matter and form. Should we at least do as much of that as we possibly can?
Granted that we can’t provide the kind of bibliographical control that would be most desirable, shouldn’t we provide as much as possible of a second-best sort? I’d argue that we should not, that you don’t compensate for a basic flaw in the instrument by making the instrument bigger and bigger. I’m talking now only about subject catalogs and subject cataloging. If it were proposed to make more of the discrete texts in a collection available by author and title, to provide more analysis but confine the analysis to bibliographical description, with no subject analysis other than what is automatically provided by title words, then I would say, by all means let’s do it. But if it comes to deeper subject cataloging, I say, forget it. In fact, go in the other direction. Abandon the idea that all the items entered in the database have to be provided with subject cataloging data; abandon the idea that the online catalog has to have the equivalent of a complete subject catalog as a subcomponent. Do subject cataloging selectively. Redefine the role of the subject catalog; rather than being a device to show everything the library has on a subject, which it never has done, think of it as having a different purpose: to serve as a readily available, convenient, universal bibliography, highly selective, confined to things on hand (not, then, expanded in the direction of being a subject index to everything available through interlibrary loan), and confined to big units—books; not a list of best books, but of books thought worth having on the premises. In some libraries of course, for example an undergraduate library in a university, it comes close to being a list of recommended books; and as a universal though shallow bibliography, a subject catalog would be no worse if it were still more selective, excluding things unlikely to be of interest except to specialists. Say that it’s universal, not in including everything you have but in excluding no topic in advance; say that it’s not meant to be complete but to provide useful starting points for one approaching a new and unknown topic, or for one who wants just one sizeable chunk of text dealing with that topic. Recognize and say that further and deeper inquiry will quickly have to abandon the subject catalog and turn elsewhere, to specialized bibliographies and indexes, and, above all, to the self-organizing apparatus of footnotes and bibliographies in books and articles. And that inquiry will then be led back, to the author and title components of the catalog, to find the particular items first discovered elsewhere. Thought of in this way, the relative shallowness of subject indexing is not a defect, and deeper indexing would not be a merit.

The upshot of all this is, as far as I’m concerned, that the unique contribution of the catalog is, after all, just what most people have always agreed it had to be, to help locate copies of books and texts that may have been learned about elsewhere. It has to be a local finding device. Its service as a convenient selective guide to the subject contents of the collection is secondary to this. It is in many ways a superficial bibliographical instrument, but this should cause no embarrassment; it is not meant to be, and should not try to be, the complete and definitive guide to the bibliographical universe, but an essential local supplement to the complex apparatus of means of discovery. In its revolutionary new online
form, that's what it will still be. But if all goes well, it will be an excitingly flexible and comfortable piece of the bibliographical apparatus.

REFERENCES

3. Ibid., p. 83 (Rule 194b).
Technology and the Online Catalog

Peter S. Graham

This paper deals with trends in technology, or perhaps it would be more accurate to say that it discusses waves; out at sea, one can point out waves and indicate that they are heading toward shore, but it is often difficult to determine which one will arrive first. The presentation assumes a lay audience in computer and communication matters; it is tutorially and taxonomically oriented. The analogy of computer system organization will be used to structure the discussion, and a controlling metaphor will be the concept of bandwidth.

Computer system organization, like Gaul, is in three parts—the central processing unit (CPU), the storage unit, and the input/output units. Regardless of computer size, from micro to NASA giant, these functions are always present. What follows is similarly organized. First I will discuss minis, main frames and their differences, and distributed processing. As storage functions I will discuss distributed databases and optical disks; and under the input/output function I will discuss long-haul networks and local networks, packet-switching, and a little bit about fiber optics and terminals.

The metaphor of bandwidth is critically important, for it deals with the issue of whether sufficient information can be transmitted within an available time. The term may be defined as describing the quantity of information transmitted per given unit of time. The origin of the term lies no doubt in the early days of radio and telephony. The wider the band of radio frequencies used to transmit the human voice, the truer the received sound will be. Most of us are familiar with the idea that the human ear, in youth, can receive sound at frequencies from about twenty to thirty cycles per second up to about 20,000 cps; and most of us know that this bandwidth decreases with age. The telephone system transmits a relatively narrow portion of the sound spectrum that we can

This paper is essentially the same as the author's presentation at the RTSD/RASD/LITA Preconference on "Prospects for the Online Catalog," on July 8, 1982, entitled "Technology Utilization in Libraries: Background, History, Current Trends in Computer Technology Availability." The author, who is head of the Book Acquisitions Department, Columbia University Libraries, wishes to acknowledge his discussions over the years on technology application with Jessica D. Gordon, John F. Knapp, and Frederic J. Meier, and the assistance of the preconference chair, Helen F. Schmierer.
hear, from about 200 to 3000 cps, which is usually perfectly intelligible but sounds rather tinny. The phone company’s saving on bandwidth allows it to pack more calls simultaneously into the same transmission line, thus effectively increasing the overall bandwidth.

There are two ways to increase the bandwidth of a communicating link: either increase the number of signal paths, or increase the signaling speed. Consider the example of a single telegraph line of the kind used on the old railroads: a single wire with the signal returning literally to ground. Morse code over this wire can be transmitted at a given rate that will define the link’s bandwidth. Getting more data over this link requires sending at a faster rate, increasing the number of signal paths on the same wire, or both. Sending several signals over the wire at different frequencies is one way of increasing the signal; laying more wires is another. Voice transmission is a third way, for it is a use of a broad band of transmission frequencies to pass information at a rate higher than dot-dash code.

Let us consider a library analogy, a library with closed stacks and one page available to enter them. A patron arrives, is served and is happy. If, however, a number of patrons arrive at more or less the same time, some will be served quickly and others will have to wait some length of time. Modern queuing theory tells us that this situation can quickly lead to loss of the library director.

One solution is to increase the speed of transmission by having the page run. A more reasonable solution is to increase the number of data paths by hiring more pages. Carrying this to its conclusion, we can open the stacks, creating as many data paths as necessary; we have radically increased the bandwidth of the system, that is, the quantity of information that can be transmitted in a unit of time. The example may seem trivial, yet it has ramifications in much of library work. For example, we must consider how many people can get at our automated catalog at one time, and how many terminals our network computer can support. The card catalog has become notorious for its defects in the past decade, but it remains a marvelously broad-bandwidth device.

One of the significant questions about the use of computers for cataloging (or for anything else) becomes: Is the necessary bandwidth available? Can enough information be transmitted in the time available? Some of the catalog use studies now in progress will provide us with necessary information on this topic, though at the moment I have observed only anecdotal evidence. We need to know more about how many people are likely to want information from a given system at a given time; and about how demand on a system is likely to rise once the initial response is seen to satisfy the initial demand quickly.

**MINIS AND MAXIS**

A few basic words about computer hardware differences should be said here. The differences among micros, minis, and maxis (or “main frames”) are relative; they are matters of speed, internal complexity, and data-access bandwidth. Table 1 illustrates these differences.

Anyone with more than a beginner’s knowledge will recognize the
TABLE 1
Differences Among Computers of Different Sizes

<table>
<thead>
<tr>
<th></th>
<th>Micro</th>
<th>Mini</th>
<th>Maxi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>ms*</td>
<td>ms, μs</td>
<td>ns</td>
</tr>
<tr>
<td>Complexity</td>
<td>low</td>
<td>low, high</td>
<td>high</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>8</td>
<td>8, 16, 32</td>
<td>32, 64, 128, or more</td>
</tr>
</tbody>
</table>

(data path in bits)

*ms = milliseconds (thousandths of a second); μs = microseconds (millionths of a second); ns = nanoseconds (billionths of a second).

oversimplifications of this figure, yet the thrust is clear: a small computer whose instruction cycle time is measured in milliseconds and which can only perform one operation at a time on a single 8-bit byte of information, is a radically different device from a main frame that executes instructions at billionths of a second on 16-byte chunks of data at a time while carrying on several operations simultaneously. The differences with respect to the requirements of your library are not always clear, and you should get a system designer (who is not your vendor) to help in making this evaluation. But you can expect a rough correlation of the following: the size of your library (however measured—collection size, circulation, number of branches); the extent of computer power required to serve your library; and the cost of the necessary hardware.

One significant problem common to almost all computers is the inherent restriction on bandwidth of a machine that proceeds stepwise through instructions. It does them one at a time. There are limits, related to the speed of light, to how fast instructions may be sequentially executed, and computers approaching that limit are in design. However, bandwidth may be increased by increasing the number of data paths. Some machines now on the market (e.g., the IBM 3081) make primitive use of two, sometimes a few more, processing units to increase overall speed; the control problems are significant enough to inhibit their full potential for general use for some time. One or two very expensive and very specialized machines have been marketed which are parallel processors, that is, their multiple units process a number of variables simultaneously using similar computation procedures. Such machines as the Cray are few in number.

Intermediate single-purpose approaches have been successful. OCLC’s development of a system using four, five, and now even more CPUs front-ended by tens of communications computers and back-ended by almost as many database processors is to my mind one of the wonders of the computing world. It is, however, unique, and has taken years to develop; and its very-long-term flexibility has to be questioned.

DISTRIBUTED SYSTEMS

What OCLC has done is related to one of the important advances in
computer systems of the past decade, the implementation of distributed systems. Distributed systems comprise two subgenres, the differences between which should be kept very clear: distributed processing and distributed databases. Both are intended to increase processing power over what can be achieved with individual systems, and both have as their purpose the location of their concern (the computing power or the data) where it will be most useful rather than at one central point.

A distributed processing system typically deals with files of data smaller than those usually encountered in libraries, files that are evanescent but need computer resources likely to be in demand. Examples include statistical computations, debugging of programs, and solution of partial differential equations. In each case there are relatively few data for all the computing power required, and the user is likely to be indifferent as to which computer in a network solves the problem. The user might decide where to send the data for processing or, in a fully distributed system, the process might be directed automatically so as to make the best use of available computing resources at a given time. Either way, the distributed processing concept is not usually our concern: library files are quite large, are relatively stable (less than one-fourth of 1 percent of a cataloging database is likely to be changed or added in a given day), the records are large and of variable length, and the processing, while quite simple, is highly interactive.

**DISTRIBUTED DATABASES**

Discussion of distributed databases brings us to the storage function of computer systems. The concept of distributing databases has to do with achieving economies and service benefits by placing the data nearer to the point of use. The most common example in library networking is probably the circulation file, whose data have in the past been of local interest only (the perception is now changing). In terms of the bandwidth metaphor, we can see that distributing the data geographically can be related to increasing the number of paths to the data.

Let me continue the earlier analogy of the library with closed stacks and a single page. A further way to increase access to library materials would be to replicate the stacks; or, in library terms, to create branches. This solution has, of course, been used, and branch library systems have been, de facto, distributed databases for many decades in all but the electronic sense. In fact, in discussing the assets and liabilities of branch library systems, one is very close to discussing the same for computerized distributed databases. Will branches have only duplicates of the central database or will they have unique copies of some titles? How does a patron know which branch has an item? Is one branch more central than others or do all have equal standing? What happens when an item is withdrawn? When the last copy is withdrawn?

To understand some of the possibilities inherent in distributing data—and some of the complexities—I think it will be useful to look at a taxonomy drawn up by Peter Enslow in *Computer.* He describes the relations that can exist between two primary components of a database:
the data itself, and the directory that catalogs the data. These components are directly analogous to the books and catalogs of a library system, and I leave it as an exercise for the reader to apply the analogy to the following seven relations:

1. A single copy of the files and the directory, maintained in local storage (the simplest case);
2. Files distributed to several locations, but only one central directory, forcing central reference before local access;
3. Complete replication: the directory and files are duplicated and maintained in toto at each location;
4. Partitioned database, with a directory and central master copy: here only part of the database is located at each outlying location, while the directory and a copy of the entire database is maintained centrally;
5. Partitioned database, with a central directory but no central master copy of the data;
6. Partitioned databases and directories, with no master file or directory (analogous to a collection of books, each with its own full index).

A seventh possibility, not included by Enslow, is appearing as desirable for library network purposes:

7. Partitioned databases and directories, with a central master copy of the total database and a master directory.

Designs for all of these, including the last, have been put forward in various computing contexts; implementation of some of them, not including the last as yet, have also appeared.

The positive aspects of distributed databases are easy to understand. Locating the data near its usage point makes sense, as tending to reduce certain communications and maintenance costs. All users of a data system will be inconvenienced if a centralized system temporarily fails, but only some will be if a portion of a distributed system fails. Presumably, a proximate database can be more quickly accessed.

But there are problems. The savings in transmitting data over shorter distances must be set against the cost of maintaining multiple databases with their own computing hardware and software. It appears that in North America this trade-off favors a centralized database—other things being equal. (This may not be the case in Europe, where the long-haul communications costs are rather higher.) Not only must a local installation of a distributed database maintain expensive equipment, it must support expensive personnel. The staff skills necessary at local data sites require specialized hardware and software knowledge of both computers and communications. These are just the kind of people who are scarce today, with resultant high salaries and high rates of turnover. It is another pressure toward centralization of databases and other functions. In this connection, it is worth recalling Barbara Markuson's useful analysis of the role of networks in librarianship as providing a locus of capital formation and of scarce skills, both of which are problems for libraries of almost all types and sizes.3

The previous problems in distributing databases can be solved, so to
speak, by throwing money at them. Technical problems remain. One is the matter of system recovery after failure. Dynamic databases are each in a state of constant change; it is important to be sure that updates in process when a system fails are done properly when it comes up again, and that they are not done twice; this is especially so when fiscal systems are involved, as they soon will be (e.g., for acquisitions systems). If databases are multiply maintained, the problem is magnified geometrically.

Multiplicity also leads to the related problems of synchronization and deadlock. To present the problem in simple terms, I will use a banking example. If Jan’s checking balance is $100 and on the same day Jan deposits $300 and Jason $100 to her account, the ending balance should be $500. The process seems simple: obtain the present balance, add the new amount to it, and store the new balance. If, however, the task performing the addition of Jan’s deposit locates the old balance ($100) just after Jason’s deposit transaction has begun, but before Jason’s is completed, the two update tasks will not take account of each other and will each store their new balance independently; depending on which does so first, the resulting new balance will be $200 or $400 (Jan will be justly annoyed and the bank will unjustly blame the problem on the computer). This sequence illustrates a problem with one of our solutions to the bandwidth problem: increasing the number of paths to the data can cause communication problems among the paths.

A solution to this synchronization problem is for the first computing task to grab control of the old balance and not allow any other task to read or write into it until the first task is completed. But in distributed systems, the solution can lead to the problem of deadlock. Let us say that Jan’s balance is maintained not only in the central bank but also in the branch bank. Let us say further that the branch computer wants to add $300 to the account, while the main computer wants to add $100 at about the same time. Each computer correctly grabs control of the old balance in its own database, so that the other computer can’t change it until the first is through. Then each computer correctly attempts to grab control of the old balance in the other computer, for the same reason. However, it finds the other old balance already locked; so the computer task (following the usual procedure) waits for it to be unlocked. In the present example, both computer tasks will wait indefinitely as neither can proceed until the other lets go; this is deadlock. A library analogy might be catalog maintenance and circulation activity on the same record, leading to a kind of bibliographic gridlock.

It is not enough to say that this kind of problem is extremely unlikely to occur; in systems where millions of instructions are executed each second, the axiom “If it can happen, it will happen” is proven daily. The examples here given, and similar problems of much higher complexity, have, of course, been solved in well-designed systems; but the solutions are costly and mean the system will be vulnerable only to more complex problems of the same type.

Overall, the problems I have described in such simple terms are so considerable that Enslow wrote, “The use and maintenance of various
organizations of distributed data bases are not yet fully understood. By this past year, one would have hoped the situation to have improved; it has, but only to a point. Holland, in May 1982, wrote:

Problems in distributed database environments frequently involve invalid updates, slow performance, excessive computer overhead, and synchronization anomalies such as deadlock. Many of these problems are still unresolved or are resolved only at great cost and overhead. Others are avoided by an overall strategic database and data management plan to control and manage a company's data.

Later on in the same article, he repeats:

Although attempts have been made to resolve these problems, several are only partially solvable or are not yet adequately resolved. From an overall standpoint, distributed database problems are so troublesome that it is desirable not to have an unconstrained distribution of data within the company.

Leonard Kleinrock, one of the best-known writers and consultants in data communications, has given his opinion that one of the failures of the computer communications industry up to this time has been its relatively poor understanding of distributed processing and distributed databases.

These are cautionary words for those looking forward to early distributed cataloging systems based on the national bibliographic utilities. Holland qualifies his warning significantly by calling for careful and extensive planning as a way of resolving at least some of the problems; yet the warnings remain. It seems to me that if the larger data-processing community is only at this stage of development, we in libraries should not expect significant movement toward distributed cataloging systems for several years at least.

And I think we should wait, as a community, until the problems in this and some other high-technology areas are well understood by our betters. We hear much about the "leading-edge" of technology; I think it is John Schroeder of RLG who has spoken of the "bleeding edge"—from experience. I prefer to speak of the blunt, trailing edge of technology. Libraries are not a noticeable market segment for any of the new communications technologies, which are being developed by industries expecting enormous profits in serving the needs of elite segments of domestic and international business. As we are not in a position to influence technological development, I think it is inadvisable for us to be early in experimenting with it. We all know the metaphor of the pioneers who have the arrows in their backs.

**Optical Disks**

Although optical disks have little to do with online catalogs, it is desirable to dispel a certain amount of confusion about them. The popular library press has dealt with the TV-compatible entertainment-oriented videodisks; the technical press discussions have been fewer and have concerned the computer-compatible digitally recorded information-storage optical disks; and neither has sufficiently distinguished the type of which it is speaking from the other.
The generic term videodisk encompasses both the entertainment-oriented videodisks and the information-oriented optical disks. In both cases, the device is a twelve- or fourteen-inch disk, recorded on both sides, which is spun at high speeds (600 to 1800 rpm) within its reading device. It has an extremely high capacity and data-transfer rate, or bandwidth. About 54,000 tracks per side allow 54,000 still pictures of moderate TV-like quality to be stored, or perhaps 2,000 pages of digitally recorded text of quite high quality. The recording technique is radically and incompatibly different for these two kinds of storage. Surface for surface, the optical disk has from ten to one hundred times the digital capacity of a magnetic disk.

The entertainment and consumer-market videodisks are of little relevance to the present discussion; they constitute a new form of material to be cataloged, and as such will place new demands on catalogers and on audiovisual-department budgets. Briefly, videodisks are characterized by TV-compatible moving images (movie, educational, or promotional material); replicatability rather like LP records; orientation toward the consumer market; read-only capability; a moderate error rate, which is tolerable as the picture quality can suffer considerably before being noticed; and relatively low cost (well under $1,000) for playback devices which in many cases can be computer controlled by, say, a home computer.

On the other hand, optical disks are a significant new storage medium that is presently worrying the microform industry a great deal and within the decade will be a great concern to the computer peripheral-storage industry. In contrast to the list above, the optical disk is characterized by digital reproductions of still images (documents, say, or book pages) or by purely digital recording of the present magnetic disk or tape kind; presently great difficulty in copying disks; one-time read-write capability on presently very expensive devices; extremely low error rates well within the margins necessary for digital storage; and costs in ranges of interest to the information industry at which it is aimed, but out of range of the individual consumer.

Optical disks are created by use of a laser beam burning small holes into a reflective surface. The beam can be precisely focused so that the holes, or pits, are on the order of a half micron in diameter in tracks spaced only two microns apart, thus allowing about 54,000 tracks on a side of the disk. The disks are read by a similar laser device which reflects a beam off the surface of the disk, reading the pits by detecting nonreflection. Some entertainment-type videodisk masters are made this way, with copies, much as LP records are by using the master to stamp out replicates. The precision requirements of information-storage disks are such that a sufficiently error-free replication process has not yet been developed; thus a copy of an optical disk takes as long to make as the original, rather like the present situation with magnetic disk information storage.

In their use as a document storage medium, the technology is based on digital reproduction of images, not on digital coding of information; that is, it is allied to techniques used in facsimile transmission.
scientific article can be stored and reproduced as it appeared on a printed page: graphs, Times Roman type, and all. In this way, while information may be digitized for disk storage, it is not searchable nor easily abstracted. Optical disks compete with microforms as an image-reproduction system, and as such its costs for very large systems are roughly comparable; the costs are going to drop further. Optical disks offer the possibility of image transmission; microforms do not. For practical purposes the transmission rate would have to be very high to be of practical use; the conjunction of this method of image recording and of local area networks (to be discussed below) is of very great interest.

The most interesting current example of library use of such a system is the Library of Congress project for storage of both print and nonprint media in lieu of the originals. LC is also using optical-disk storage of catalog card images in its catalog distribution service.

In the longer run, falling costs of optical disks will allow their use as digital-storage media rather like present magnetic-disk packs, only in more compact form; this eventuality will be hastened when optical disks can be copied more readily than now. In any case, I hope it is clear that video and optical disks are of no direct relevance in present planning for an online catalog: they are passive storage devices and as such will simply offer another tactical option for system implementers.

**COMPUTER NETWORKS AND COMMUNICATIONS**

With the discussion of local area networks, we move to the final portion of the computer analogy with which we began: the input/output structure. It will be useful first to review larger-scale networks and how computers communicate data, both internally and externally.

Computers communicate with each other by means of networks. In fact, the interior structure of a computer is a miniature network, in that the main components (the central processing unit (CPU), the storage unit, and the input/output units) also need to pass data back and forth under some kind of control. The distinction between the types of networks available to designers for use inside and outside a computer has become more and more fuzzy.

Network descriptions usually begin with topology, that is, with the distribution of network nodes in space and their interconnection. Some types are more familiar than others, but all of these are now in use.

1. Unconstrained: in this model, each node is a computer that is probably itself the center of a star network (see below). Each CPU must accommodate the others, which may be of different manufacture, by providing software that can receive and send messages to possibly quite differing systems. Message routing from one CPU to another is decided by each node. The resulting network efficiency is thus fairly low, particularly for short, frequent messages. Failure of a node in such a network is only of moderate significance to network activity as a whole. This type of network is like that which was occasionally proposed for the national library network in the late 1970s.

2. Ring: this model is seldom seen in long-haul (geographically large) network design, though it has been used for local area networks...
and within computers. In such a model messages in a predetermined format are passed around the ring from one node to the next until they reach their destination. Typically all nodes use the same programming and message structure, a method that simplifies and speeds up transmission. Failure of a single node is catastrophic to the network, though various backup solutions are often implemented (redundant alternate-node connections, reverse-direction pathways, and the like) with the usual concomitant cost and complexity.

3. Bus: in this model all nodes are connected to a single long communications link, or bus. While common in minicomputer design, it is relatively new in intercomputer communications. A single node failure affects only devices attached to that node, and a break in the bus only separates the network into two sections. When a message is placed on the bus by a node, it is received by all but must be ignored by all but the receiving node. The control mechanisms for determining which node places a message on the bus are moderately complex.

4. Star: by far the most familiar network form; the library utilities, with a central computer supporting terminals on communications lines, are archetypal examples. The failure of a given communication line or terminal is only locally a problem, a fact we all know as well as we know that the failure of the central computer is catastrophic to all parts of the network. Unlike the other network forms, routing decisions take place at the center, increasing its relative importance.

Communication within a computer takes place differently from that between computers. A computer sends an 8-bit byte to one of its components over eight separate lines (knowledgeable readers will recognize that I am ignoring parity bits and other control functions). Within the short distances found within computers, we have confidence that the eight signals will all arrive at their destination at the same time, allowing the next byte then to be sent. This is not true over longer distances, owing to variables such as switching delays, circuit resistance, differential cable lengths, and the like. The situation is like eight runners at the starting gate; a moment after the gun goes off, a narrow-angle camera a few feet from the start can take a picture across the lanes which will capture all eight of the runners. A similar camera at the 100-yard finish will be fortunate if half the runners are in view. The displacement is termed "skew."

Sending the 8 bits at once is described as parallel transmission (and in many computers 16, 32, 64 or even 128 bits are sent in parallel simultaneously). Between computers, taking account of the longer distances, information is sent serially, one bit after another. It is as if the runners lined up in packs of eight and ran through the finish line one at a time.

Until recently, serial transmission has been slow in comparison with internal computer speeds. The conversion from parallel data to serial, and the reconversion, has required various solutions; the use of external lines of various quality owned by third parties (mostly the phone
company) has required slower speeds and accommodations, including the development of the modem; and the terminals at the other end of most lines have been slow devices anyway, requiring the computer end to restrain its transmission and otherwise to occupy itself in the interims. It is prohibitive to lease a communication line from New York to each terminal connected to the system in Chicago, so a whole industry has grown around techniques of sharing communication lines. Polling, terminal addresses, multiplexing, bi-synch protocols, and front-end communications controllers are all terms for devices and procedures necessary to handle computer communications over multi-drop lines—increasing their bandwidth even though the actual transmission speed has remained low.

The need to optimize communications among large numbers of nodes on a relatively slow communications network, moderately liable to failure, has led to the important concept of the message packet; and packet-switching is now at the heart of both long-haul and local area network development. Real-world messages tend to vary greatly in length, in sources, and in destinations. A packet-switching network breaks up all messages into standardized packets of predictable length, and assigns sequence numbers to them; requires standard addresses for destination and origin that are appropriately plugged into each packet; and requires standardized packet-switching programming at each network node. This programming allows the node to forward packets to the next node on the way to the destination with a minimum of overhead; just as important, it allows an alternative route easily to be chosen if a node on the more direct route is known to be busy or to have failed.

The result of network development through the 1970s has been that the large expensive computers and intercomputer communications have been well served. Wide-bandwidth channels using packet-switching technology at relatively high speeds of 56,000 bits per second and more ("fat channels") have achieved economies of scale. Traffic is high on networks such as the ARPAnet; the computing resources available are very great; and the result has been high efficiency in terms both of costs and response time at the computer.

The problem, as Leonard Kleinrock has pointed out, is that the users at the periphery are not yet well served. Individual terminals, by the nature of the human interface, are low-speed and low-traffic devices, and therefore inefficient. When I am at a terminal, Kleinrock says, communications analysts have the following problems:

- they don’t know when I at the terminal will require resources;
- they don’t know how much resource I will need when I do;
- mostly, I won’t need any resource (I’m scratching my head);
- when I do, I want it now. Inefficiency is built in.

**LOCAL AREA NETWORKS**

A solution may lie in local area network developments which are on a front of extraordinarily great activity and must now begin to affect our
Local area networks (LAN) are characterized by the following:

- extremely high data transmission rates, or bandwidth, ranging from 1-12 Mbits per second (1 Mbit, or megabit, is 1 million bits);
- compatibility between all devices attached to the network, regardless of speed, type, or manufacturer;
- use of relatively inexpensive transmission media, whether twisted-pair telephone wire, coaxial cable, or fiber optics—in most cases, the inexpensiveness deriving from its not belonging to the phone company;
- topology within a limited geographical area: a campus, a small town, a city-hall complex;
- connection with other networks (long-haul or LAN) through standardized "gateways";
- no central control point, and therefore resistance to failure.

LAN design calls for the network structure itself to take over intercommunication functions. The attached devices, whether computers or terminals or other devices (such as, say, an optical-disk storage device for facsimile document retrieval) present the network with a message and address. The development is a mixture of hardware developments (which have resulted in the high speed and the switching capabilities) and software developments (the compatibility and packet programming). Manufacturers and promoters are now talking of wiring buildings, to say nothing of offices, to allow easy connection and disconnection of office equipment to a proprietary local network. Many university computing authorities are planning in much the same way. The high speed, ease of connection, and lack of control center spell the end of the hegemony of the traditional star network in the local situation; I hope it is clear that there are very great implications for the technical planners of online catalogs over the next decade.

As an aid to those of us who are not the technical planners but who must at all costs look over their shoulders, I want to sketch out some of the terminology and major tendencies in LAN development that are now emerging. As usual, network discussions begin with topology; and two of the familiar forms have emerged preeminent in LAN development: the ring and the bus.

The ring, as previously described, allows for passage of a message around a circuit until it reaches its destination. In the LAN implementation, the network is an active component; the nodes are repeater stations, either passing the message on or stripping it off if this node is the addressee, and passing it to the attached device. As messages are a predetermined length, each fits into a "slot" of time available for transmission from one node to the next. Some ring networks allow only one message at a time to go around the circuit, but others allow more. All networks must have a control mechanism, allowing nodes to know when they may transmit; ring LANs tend to use the "token-passing" technique, in which a specific character-space in each message packet indicates whether the slot may be used. If the bit pattern is not that of
the token, the node may change it to the token and place its message into
the remainder of the packet. IBM is one vendor that has indicated a
preference for the token-passing ring LAN technique. The bus topology is the other form used in LAN development so far. The bus is mechanically simpler than the ring, at a cost of requiring somewhat more complex software techniques, because the bus is a passive scheme, relying on the nodes to place messages on the network in an orderly fashion. An electrical signal placed on the bus will be available to all nodes on the bus more or less simultaneously; it is up to the correct node to pick up its own address and receive the message, and it is up to the other nodes to ignore messages not addressed to them. In this way the nodes resemble terminals on a multidrop communication line that must respond to their own polling address and to no other, though they can "hear" the messages to and from all other terminals on the line.

In a bus network it is entirely possible for two stations to begin transmitting simultaneously; but as both signals propagate along the bus, they would intermingle chaotically. One of the properties of LANs is that there is no central unit for adjudicating priorities, as there is in a star network which polls and addresses terminals in a controlling fashion. The control technique of a bus LAN requires each node to listen before and during its transmission. If the node is receiving, it must not transmit; and if it begins receiving in the first few bit-spaces after it begins sending (a "collision"), it must stop and begin the listen-before-sending cycle again. As each node waits a random number of milliseconds before starting the cycle again, a second collision is unlikely. The procedure has come to be known as "carrier sense multiple access with collision detection," and the clumsy term CSMA/CD has now come to be part of data-processing jargon. The Xerox-inspired "Ethernet" is one major development using these techniques.

The two techniques are wholly incompatible with each other and both are being developed by substantial market providers, ensuring their continued existence for some time. Gateways are possible between these incompatible network forms, as they are between LANs and the long-haul networks.

**Transmission Media**

The choice of LAN type is not constrained by transmission medium. Three types have been proposed so far. The first is twisted pair, which is telephonelike wiring (though not owned by the phone company), inexpensive, familiar, and easy to install. Speeds of up to 10 Mbits/second can be achieved over fairly short distances, and splicing—adding new connections—is an easy matter. The possible speed drops rather quickly with distance. Twisted pair is also somewhat sensitive to electrical or radio interference, making it inconvenient in certain heavy industrial or military applications (one reads in the literature of the need to protect communication systems from intense radiation bursts). IBM’s initial LAN installations will probably be with
twisted-pair wire. Migration to higher-quality connections at a later date will not be difficult.

Coaxial cable is another medium of choice. Physically the cable is identical to that used by cable TV systems; it is a single copper wire at the center of an insulating covering, which is further covered by a braided shielding wire with the same axis, or center, as the copper strand—thus coaxial. This cable is less subject to interference and is capable of high-speed transmission over longer distances than twisted pair. Splicing is somewhat more difficult. Its great advantage lies in its facility for transmission of several signals at once. With the use of what are called baseband techniques, a single signal can be sent at rates up to 10 Mbits per second. With the use of broadband techniques, several digital signals of 1 Mbit/second can be transmitted simultaneously with a wide-bandwidth video signal, a method that allows for considerable flexibility in the use of the cable system. Both techniques are being implemented by LAN vendors.

Fiber optics systems are being developed and will see use with LANs. In such a transmission line, a light wave is sent along the inside of a thin strand of glass. The strand has one refractive index at its core, and another slightly higher around its perimeter, forcing the light signal to propagate along its axis. Signaling rates are expected to be at rates of about 1M to 10M bytes per second, paralleling the internal speeds of computers attached to the network. It is not at all sensitive to electronic interference. It is presently rather difficult to splice into the line to add new connections, but is about the same size and as easy to install as the other two forms of cable. When this high-bandwidth technology is mature—within the decade—it will offer enormous speed and reliability advantages.

**TERMINALS**

Let us accept the many criticisms that have been made of the card catalog; the fact remains that it is still a tool in which a great deal of information can be seen in a very short time. Recall for yourself the experience of flipping through a set of cards looking for the one you want; under a populated heading it is easy to flip through a dozen or twenty cards, scanning them for important points (imprint date? subject heading? editor’s name?) almost as quickly as you can flip them. Card catalog bandwidth is high because of the high capacity of human perception.

In this respect I am disappointed in what current CRT terminals offer the library user; in spite of the powerful Boolean search capabilities now provided, present terminals now actually inhibit access. Their screens are small: twenty-five lines by eighty characters is inadequate. Terminals presently lack data storage space of any consequence—by which I mean many thousands of characters, so that it can take over some of the interactive responsibility. Response time is slow; we correctly speak of eight-second response time as a goal for 95 percent of our transactions, but we would be infuriated by being able to move cards in a catalog at such a slow rate.
Finally, in typographical terms the modern CRT is positively a barrier. Everything we have learned about letterforms and readability over the centuries—from Dürer to Baskerville and the twentieth-century masters like Edward Johnston and Stanley Morison—tells us that distinctions in size, boldness, placement, and spacing of letters are of the greatest importance in assisting the reader to distinguish what is being read. The computing industry has spent billions on terminals over the past decades, but has ignored this issue. The use of color may help in the future, but at present, to compare an LC printed card with the screen or printed output of a bibliographic utility is an exercise in sadness:

Look here upon this picture, and on this,  
The counterfeit presentment of two brothers.  
See what a grace was seated on this brow...  

The recent Hildreth report from OCLC on the user interface of online catalogs spends some pages discussing the different output screens of a dozen online cataloging systems, and makes a number of useful evaluations; while worthy, the distinctions are in a way without a difference. Through no fault of the system designers—they must use what is available—none of the screens can be called genuinely helpful. Library systems are in danger of giving new meaning to the phrase “terminal illness.”

CONCLUSION

One of the important goals in librarianship today is increasing the bandwidth: increasing the information flow with speedier service and using an increasing number of means and media, whether for children, scholars, inner-city patrons, or suburban readers. I have argued that the same is true for the technical implementation of the online catalog, and my argument has led me to certain tentative conclusions, liable to change as the technical and political situation changes.

One of our present problems is the limit on speed inherent in any computer. It is interesting that the solution—some kind of distributed system—is philosophically one that libraries have been articulating for some time. Sharing of resources is now widely recognized not as an ideal but as a necessity, and it is becoming so in computing terms as well as for library materials. In the long run, a star network won’t work, whether one is speaking of library networks or computing networks, for while there are inherent limits to a single processor or library, there are no inherent limits to demand. No library can stand alone any more, whether we speak of books or of computers.

I have tried to indicate where technology may provide solutions to our problems, but also to indicate that many of the solutions are not yet in place. There are bibliographic networks, yes; and technical processing systems of great usefulness. Are there online catalogs? There are a very few systems that might be called catalogs and not just computerized databases.

It would be helpful if the profession, including its computer
technicians, would review why Cutter’s definitions of a catalog have served us so well for most of a century. Must something fit the definition before we use its name or, like Adam, can we name the beast simply because we have the arbitrary power to do so? At conferences, and in journals and at staff meetings, we tend to reify “the online catalog” by speaking of it. It doesn’t exist, but we keep talking as if it does, gradually becoming honestly persuaded that it does exist—or after all, how could we have been talking about it for so many years? And by exist, I don’t mean in isolated examples, but in the generality, like cars as opposed to spacecraft.

I am recommending a retreat to realism in the debate I know is now taking place in many of the nation’s libraries. Few libraries today can realistically decide whether or not to have an online catalog, for they really aren’t being offered one; they can however decide to choose system A over system B. A library might not choose even an already-existing system, but someone’s prototype system (maybe its own)—in which case it is asking for trouble. Some of us may be willing to make this choice, and it is due to this willingness that some leading librarians and libraries have led the profession forward at times. But we have a responsibility in such cases not to fool ourselves or our management that we are choosing between a Volkswagen and a Cadillac; we’re buying a new Boeing 767, and only the prototypes of those have flown so far.

What are the more cautious of us to do? There are some simple but solid guidelines. First of all, work with others; don’t go it alone. Particularly for libraries, our fiscal and intellectual resources will be better used when pooled.

Second, keep the data clean. To paraphrase Dr. Johnson, an institution that does not mind its data will not mind much of anything. This may seem a truism, as for decades we’ve all minded our card catalogs and circulation files; yet there are many libraries who have paid good money for years of archive tapes and have done nothing but stack them. And there are librarians who speak of better, or shorter, or more cost-effective, or more locally useful forms of bibliographic data as if they had forgotten what library interdependence means. I will invert a maxim that gained currency a few years ago: in terms of our technological tools, the future is shorter than the past. The mechanisms for creating, displaying, and distributing catalog data are changing constantly. Yesterday libraries spoke of having their own computer; today few do; tomorrow, the minicomputer revolution may arrive. In a startlingly short time microforms may be outmoded (did you expect to hear that five years ago?). Network types, communication protocols, display devices are being planned at a rate which Kleinrock has said makes data communication systems obsolete before they can be brought to market. The cautious approach will recognize that the data itself is what is remaining constant: systems come and go, but data abides.

Third, we should therefore use technologies that have arrived at some maturity, or which are fully supported by industries capable of dealing with the setbacks of new system implementation. And fourth, we
should work to influence technological development where we do it best:
in improved access techniques (subject headings, for example), and
above all in matters of authority control. Such matters are uniquely our
past, and they will serve our future well.

REFERENCES

1. Olin H. Bray, Distributed Database Management Systems (Lexington, Mass.: Heath,
1982) is a very good general text.
2. Peter H. Enslow, "What Is a 'Distributed' Data Processing System?" Computer
5. Robert Holland, "Distributed Databases: Decisions and Implementations," Data
Communications 11:97 (May 1982).
6. Ibid., p.106.
7. Leonard Kleinrock, "Computer Communications," lecture at Columbia
University, April 21, 1982. Further cautions are expressed by Hector
Garcia-Molina and Daniel Barbara, "The Cost of Data Replication, " in Seventh
Data Communications Symposium: Proceedings ([N.Y.]: IEEE Computer Society Press,
1981 [LC 81-82744]), p.193-98. The entire volume of proceedings (hereinafter
"7th DCS") is current and useful.
8. A good but technical report is R. Barrett, Developments in Optical Disk Technology
and the Implications for Information Storage and Retrieval, BL R&D Report 5623 (Boston Spa:
useful material appears in Efrem Sigel, Mark Schubin, and Paul F. Merrill, eds.,
Video Disks: The Technology, the Applications and the Future (White Plains, N.Y.:
10. For example the ADONIS document delivery project now in development by six
scientific publishers, including Elsevier, from whom information may be obtained.
See also: Jacque-Lynne Schulman, "Video PATSEARCH: A Mixed-Media
11. Library of Congress, "Request for Proposal: Optical Disk Based Image
Preservation and Retrieval System: Pilot Projects for Print and Non-Print
Materials" (Washington, D.C.: Library of Congress, April 12, 1982),
mimeographed.
12. For recent comments on relative costs, see Velina Houston, "Optical Disks in
Spotlight, but Tag Seen Still Too High," MIS Week 3:1 (May 12, 1982).
14. For further information see the special issue on packet communication networks of
16. The best brief overview, very current, is William P. Lidinsky, "Insights into the
Implementation and Application of Heterogenous Local Area Networks," in 7th
DCS, p.52-62. David D. Clark, Kenneth T. Pogran, and David P. Reed, "An
Introduction to Local Area Networks," Proceedings of the IEEE 66:1497-1517 (Nov.
1978) is still the important starting point. Recent and useful is Ronald W. Gibson,
"A Primer for Evaluating and Purchasing Today's Local Networks," Data
Communications 11:147-56 (June 1982).
7th DCS, p.149-55; and C. Kenneth Miller and David M. Thompson, "Making a


19. An analysis of one functioning LAN has shown that in fact less than 1 percent of the messages have to wait, and less than .1 percent are involved in collisions; and that the line is not in use at all over 95 percent of the time: John F. Schoch and Jon A. Hupp, "Measured Performance of an Ethernet Local Network," *Communications of the ACM* 23:711–20 (Dec. 1980); there is a useful summary review of this article in *Computing Reviews* 22:315–16 (July 1981).


29. Essential journals for monitoring the marketplace are *Datamation* and *Data Communications*, both monthly; *Computerworld*, a weekly; and *Information Technology and Libraries*, quarterly.
The functional characteristics of twenty online public catalogs in libraries in the United States are described. Database and technical characteristics, bibliographic questions, and costs also are briefly noted.

The world of online catalogs is exciting, partly because of the new and enhanced capabilities for bibliographic access that the catalogs make possible, and partly because the characteristics of these catalogs are constantly changing. Some have been in existence for several years, others are still in the planning stage, and all of them are still under development. This mutability makes things interesting to observers as well as to participants in the development process, but it also makes it very difficult to describe and compare the systems. Since none of the systems is really a finished product, it is difficult even to decide which systems to describe and compare. For the purposes of this paper, the only systems included are those that are operational, at least in a prototype mode, so that a functioning "thing" could be described, and the discussion has been limited to systems in the United States. Twenty systems, selected by these criteria, are described, listed in table 1 by the name of the institution that developed each.

Excluded, because they are not yet operational, are systems at Missouri and Illinois, to be based on software from the Washington Library Network, and a system being developed jointly by the University of Wisconsin and the University of Chicago in cooperation with IBM. The Online Computer Library Center (OCLC) and the Research Libraries Information Network (RLIN) are also excluded, because they were not designed primarily for public access, even though some libraries are using them that way. A few other systems have not been included simply because the institutions where they are located did not reply to the letter questionnaire on which much of this paper is based. Any other omissions are inadvertent, and I would appreciate having such omissions called to my attention.

There are two other caveats. Because of the rapid changes in the field, some of the information in this paper may be inaccurate, or may soon

Stephen R. Salmon is assistant vice-president for library plans and policies in the University of California Systemwide Administration. He presented an earlier version of this paper at the RTSD/RASD/LITA Preconference on "Prospects for the Online Catalog," on July 8, 1982.
## TABLE 1

**ORIGIN AND SCOPE OF TWENTY ONLINE PUBLIC CATALOGS**

<table>
<thead>
<tr>
<th>Organization</th>
<th>Online Catalog</th>
<th>System Name</th>
<th>Origin</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avatar Systems</td>
<td>ILS</td>
<td>x*</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Beth Israel Hospital</td>
<td>PaperChase</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Biblio-Techniques</td>
<td>BLIS</td>
<td>x*</td>
<td>x†</td>
<td></td>
</tr>
<tr>
<td>University of California</td>
<td>MELVYL</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Carlyle Systems</td>
<td>TOMUS</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Claremont Colleges</td>
<td>TLS</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>CLSI</td>
<td>PAC</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Dallas Public Library</td>
<td>LSCAN</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Dartmouth College</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>DataPhase Systems</td>
<td>Inquiry</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Geac International</td>
<td>Public Query</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Inforonics</td>
<td>DB-SEARCH</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Library of Congress</td>
<td>SCORPIO/MUMS</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mankato State University</td>
<td>PALS</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Northwestern University</td>
<td>LUIS</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Ohio State University</td>
<td>LCS</td>
<td>x†</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Pikes Peak Library District</td>
<td>Maggie’s Place</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Universal Library Systems</td>
<td>ULISYS</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Washington State University</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Washington University School</td>
<td>BACS</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>of Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Developed by libraries and then licensed to commercial organizations that now distribute them.
† Developed by a commercial organization for a library.
‡ Capability exists.

The names of the twenty systems, also shown in table 1, include not only anthropomorphic names such as “LUIS,” “MELVYL,” “TOMUS,” and “ULISYS,” but also initialisms such as “ILS,” “TLS,” and “LCS”; descriptive terms such as “Public Query” and “Inquiry”; names more reminiscent of restaurants or TV shows, such as “Maggie’s Place” and “PaperChase.” Two of the systems have no names.

About half were developed by libraries and the other half by commercial companies. Actually, the line is a fuzzy one; in two cases, the systems were developed by library institutions and then licensed to commercial organizations that now distribute them, and in one case the system was developed by a commercial organization for the library that now distributes it. Most were designed for single institutions, but nine of them serve as online union catalogs, a fact that makes them both more useful and more complex.
The balance of this paper discusses specific characteristics of individual systems; first, the functional characteristics, including a consideration of how the systems get the user started, how they search for bibliographic records, how they display the records, and how they assist the user when help is needed; second, a brief mention of some bibliographic questions; then a bit of information on database characteristics and technical characteristics; and finally, a word about costs.

**FUNCTIONAL CHARACTERISTICS**

**Search Method**

To begin the search for records, most systems take one of two approaches. The first is generally called the *menu* approach, meaning that the user is given a list of options at each step of the search, and chooses the option wanted by pressing a number or letter on the keyboard. The other method is usually called the *command* approach, meaning that the user types in words, letters, or symbols to tell the computer what to do.

*Menu Approach.* Here is an example of a simple menu approach.

**CHOOSE THE TYPE OF SEARCH YOU WISH TO PERFORM**

1—BY AUTHOR  
2—BY TITLE  
3—BY SUBJECT  
4—OTHER  
OR  
ENTER A TERM

The instructions are brief, and the user merely picks a number to get started. Here is a similar approach, except that the user picks a letter.

**Function: Look up a book**

Enter the letter that corresponds to what you want, then press ‘SEND.’

T— Look for a book using a TITLE.  
A— Look for a book using an AUTHOR.  
C— Look for a book using a CALL NUMBER.  
S— Look for a book using a SUBJECT.  
X— Go back to Main selection menu.

ENTER:

From the standpoint of human engineering, or *ergonomics* (as it’s coming to be called), the use of the alphabet is better, because there’s a mnemonic value to the letters.

Some menu approaches are more complicated. In using the Biblio-Techniques “master menu” (shown in figure 1), four steps are necessary. The user must first “tab” the cursor on the screen to certain spaces, and fill in the text to be used in the search; then check the “corporate and conference words indicator,” and if necessary tab the cursor to that position and type “yes” or “no” to change the indicator.
Characteristics of Online Public Catalogs

Bibliotechniques Library & Information System (BLIS)
Easy Access Searching - Master Menu

(FILL IN INFORMATION)

Title Words:
Author:
Subject:
Subject:
Series:

( - Corporate & Conference Words Search

( - Library Catalog
TAB to
( - Index of author, subject, series terms & cross references
arrows
( - Branch location & call number
( - Network libraries locations & call numbers
Use ENTER
to search
( - Easy Access Searching Guide
( - Library Information

Easy Access Instructions
1) Use the TAB key to move the cursor
2) Fill in information, use ENTER key to search
3) Use the CLEAR key to refresh the Master Menu
4) Use a "*" at the end of a non-specific search term,
e.g., Author: Brown, John *

Figure 1

value; then tab to an arrow to indicate what type of information is to be retrieved; then press the "ENTER" key. As with several systems, this menu can be modified and simplified if desired, by any library using the system.

In the California system, the user is first greeted by the introductory screen shown in figure 2. A choice must then be made between the menu

WELCOME TO MELVYL

The University of California's Prototype On-line Catalog

MELVYL contains 733,412 records representing approximately 1.3 million books in the University of California libraries.

*****************************************************************************
* This is NOT a complete record of books in the UC libraries.  *
* Check your local catalog if you don't find what you need in MELVYL.*
*****************************************************************************

** Type the number of the search mode you want, then press RETURN.
Or, for more information, type HELP and press RETURN.
1. LOOKUP mode (for new or occasional users)
2. COMMAND mode (for users familiar with COMMAND language)

Figure 2
First MELVYL Screen
Welcome to Lookup Mode.

You can search for books at any one of the following campuses or groups of libraries. Select one of the following, and type its code.

- UCD (Davis)
- UCI (Irvine)
- UCR (Riverside)
- OCSF (San Francisco)
- UCB (Berkeley)
- UCSC (Santa Cruz)
- UCU (Los Angeles)
- UCSD (San Diego)
- OCSB (Santa Barbara)

You may search for books in one of two ways:

- by AUTHOR and/or TITLE
- or by SUBJECT

Type the number you want below or type HELP, then press RETURN.

1. AUTHOR / TITLE search.
2. SUBJECT search.

The user must thus go through three screens before beginning to search with the menu approach.

Command Approach. In the command approach, some systems have no introductory screens or explanations. The user must learn the commands elsewhere (from a manual, or a sign, or a library staff member, or a friend) then start entering commands. Others give the user a little help with an introductory screen of explanations and an option to get further instructions. An example from the Mankato system is shown in figure 4.

The instructions can sometimes be made clearer by including examples of commands and a brief definition of the system or its scope. The introductory screen in figure 5, taken from the Carlyle system, is an illustration. It actually has four components: it explains what the system is; it explains how to use it (i.e., by typing commands); it gives examples of valid commands; and it tells how to get more help.

The dividing line between the menu approach and the command approach can be fuzzy. With the introductory screen from the Dartmouth system, the user chooses a function by typing what looks like a command.
The Online Catalog will be down Aug. 23-25 for hardware maintenance and installation of new equipment on the UNIVAC-1100 system. The Online Catalog may also be down Aug. 26 for loading of the new Online Catalog System.

Type HELP at any time to obtain assistance.

**THE BASIC SEARCH COMMAND FORMS ARE:**

- **AU** Last-name First-name Middle-initial
- **CO** Author-last-name First-title-word
- **SU** First 4 Subject Reading Words

*where (BO) is a boolean operator (AND NOT OR)*

Send messages to the computer by DEPRESSING the (NEW LINE) key. Use the (BACK SPACE) key to BACK UP and type over mistakes.

Other search options are available - type HELP.

---

**Figure 4**
First PALS Screen

**Figure 5**
First TOMUS Screen

Welcome to the Dartmouth Libraries On-Line Catalog!
(To correct typing errors, use the BACKSPACE key in the upper right)
There is a printed demonstration of this on-line catalog on or near this terminal.
To start searching for a book or information, type SEARCH now.
And be sure to press RETURN after you have responded to this and every question.
Search, Browse, List, Explain, Help, Goodbye then press RETURN?

It’s actually a menu system, however, because the user is guided at each step by the computer, and given a list of options from which to choose. The Northwestern system, on the other hand, starts out looking like a menu system (see figure 6). After a selection is made from the introductory menu, however, the system displays instructions on how to use the commands, and from there on it’s a command system.

A unique approach is used by the PaperChase system at Beth Israel
LUIS: LIBRARY USER INFORMATION SERVICE

LUIS can be used to find BIBLIOGRAPHIC, CALL NUMBER and LOCATION information for materials held by Northwestern University Libraries and by the Garrett and Seabury-Western seminary Libraries. (Use the card catalog for materials not in the LUIS database.) CIRCULATION information for materials charged out through the computerized system is also available.

TYPES OF SEARCHES:
FOR INTRODUCTORY SCREEN FOR TITLE SEARCHES:
AUTHOR SEARCHES:
SUBJECT SEARCHES:
FOR USERS ALREADY FAMILIAR WITH LUIS:
FOR CIRCULATION INFORMATION SCREEN:

CALL NUMBER MUST BE KNOWN

TO CORRECT A MISTAKE, type over the error or press CLEAR to start over.
TYPE COMMAND AND PRESS ENTER

Figure 6
First LUIS Screen

Hospital in Boston. The first display merely says, "Look for": The user types a term or a string of letters, and the system looks for that string in all the indexes. If it appears in more than one—perhaps in the title index and the subject index, or in some cases the author index as well—the system asks which category is wanted, and from then on it’s a menu system.

Access Points

The access points, or search elements, of each system are shown in table 2. They include author, title, and subject in all systems. Most include corporate authors and conference names in the author index; in four systems, however, these must be searched in a separate index. Uniform title is usually included in the title index, but is searched separately in one system. Series title is usually in the title index, but is separate in five systems; in the Claremont system, it must also have at least the first parenthesis. Fifteen systems allow searching by call number; thirteen by LC card number; eleven by ISBN and/or ISSN; five by OCLC number; two by CODEN number; and three by document number.

After the search element is chosen, the response in the menu approach varies from system to system. The most common response, found in five systems, is a prompt of some kind, such as an arrow, or "author" followed by a question mark, or "enter author," or "please enter author." One system gives no prompt, and requires that the search element be typed immediately after the letter representing the search element; a subject search, for example, is made thus:

SSOLAR ENERGY
### TABLE 2
**Access Points Provided by Twenty Online Public Catalogs**

<table>
<thead>
<tr>
<th>Online Catalog</th>
<th>Aut</th>
<th>Tit</th>
<th>Sub</th>
<th>Cor</th>
<th>UnT</th>
<th>Ser</th>
<th>CaN</th>
<th>CdN</th>
<th>StN</th>
<th>OcN</th>
<th>Coden</th>
<th>DoN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avatar Systems</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x†</td>
<td>x†</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Beth Israel Hospital</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biblio-Techniques</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>University of California</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carlyle Systems</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claremont Colleges</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x†</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLSI</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dallas Public Library</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dartmouth College</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DataPhase Systems</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Geac International</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x†</td>
</tr>
<tr>
<td>Infornics</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x†</td>
</tr>
<tr>
<td>Library of Congress</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x†</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Mankato State University</td>
<td>x</td>
<td>x</td>
<td>x†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northwestern University</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Ohio State University</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Pikes Peak Library District</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal Library Systems</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington State University</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Washington University School</td>
<td>x</td>
<td>x</td>
<td>x†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x†</td>
</tr>
</tbody>
</table>

* Aut = author; Tit = title; Sub = subject; Cor = corporate heading; UnT = uniform title; Ser = series title; CaN = call number; CdN = Library of Congress card number; StN = International Standard Book Number or International Standard Serial Number; OcN = OCLC number; Coden = CODEN number; DoN = document number.
† Accessed through separate index
‡ Available to staff only.

Some systems provide formatting instructions, especially those that require the author’s last name first, followed by a comma, and so forth.

**ENTER AUTHOR (AS MUCH AS KNOWN) AT LEAST 3 CHARACTERS. INCLUDE COMMAS, SPACES, PERIODS ONLY IF PRESENT IN THE RECORD: LAST NAME FIRST (EXAMPLE: HENLE, JAMES M.).**

One system provides a “work form” to assist in the formatting.

**PUBLIC CATALOG AUTHOR SEARCH**

**LAST NAME**  **FIRST NAME**  **MIDDLE NAME**

An unusual approach is taken by the CLSI public access catalog. It employs a “touch” terminal, with touch-sensitive pads behind various parts of the screen. There is no keyboard. This method means that everything must be chosen from a menu, and the contents of the search element—the author’s name, or the words of the title or subject—cannot be keyed in. Instead, the alphabet must be successively partitioned by touching the screen until the precise entry wanted is reached. Figure 7 illustrates the steps necessary to find the title *Off the Wall* in the CLSI system as installed at the Evanston Public Library. The user first touches “Title”; then in each succeeding display the user must touch the title that alphabetically precedes the desired title, namely *My Last Book*, then *Nuclear Proliferation*, then *Off-Season Football Training* (if the user
knows the filing rules well enough to know that "Off-Season Football" comes before "Off the Wall"), then Off the Record with F.D.R., and finally Off the Wall. The system then presents a choice of four specific entries.

HELP

Welcome to Evanston Public Library's new Public Access Catalog (PAC)

This is a touch sensitive terminal which provides up-to-date listings of the library's holdings by author and title. Subject access will be gradually added during the year.

We hope that you will enjoy using it.
Librarians will be happy to assist you.

Author
Title
Subject

START
OVER

HELP
CL SYSTEMS

"And Other Poets,"
Beginner's Gde To The Skies
Collected Poems of W. Stevens
Edu For Admn Careers in Gov
Gente Senza Storia
I like Old Clothes
Lincoln Kangaroos
My Last Book
Plant Production Control
Sea And The Stone
Taxation, USA

START
OVER

HELP

Figure 7
Searching Off the Wall in the PAC System: Displays 1 and 2
My Last Book
Nat l Dctry of Arts Supp. By
New Federal Wage-Hour...laws
Nine Great Plays
Nuclear Proliferation
Old Sheet Music
Operational Philosophy
Outdoor Rambles
Parenting The Difficult Child
Pennington’s Heir
Photography Afield

START
OVER

Nuclear Proliferation
Nursing Mother
O.K., Historian Without An...
Obstacle Race
Ozeretz
Of Cats & Men
Off-Season Football Training
Oft-Told Tales of Lincoln
Oklahoma
Old English Church
Old Magic

START
OVER

Nuclear Proliferation
O.K., Historian Without An...
Obstacle Race
Ozeretz
Of Cats & Men
Off-Season Football Training
Oft-Told Tales of Lincoln
Oklahoma
Old English Church
Old Magic

Figure 7 (Continued)
Searching Off the Wall in the PAC System: Displays 3 and 4
Off-Season Football Training
Off The Record With F.D.R.
Off With Their Heads
Offensive Traveller
Office Bldg Design
"...Officer Down, Code Three"
Official Am Ski Technique
Official Chess Rule Book
Official Illustrated Guide
Official Souvenir Program
Offl Maj. Leag. baseball Bk

---BROUSE REVERSE---
---BROUSE FORWARD---
DISPLAY ALL
HELP

START OVER FAST SLOW FAST

FROM:
Off The Record With F.D.R.
Off-The-Road Locomotion
Off The Rocks
Off The Wall
Off To A Good Start
Off To College
Off To Mexico
Off To The Right Start In...
Off Washington Square
Off-Wheel Pottery Book

TO:
Off With His Head

Figure 7 (Continued)
Searching off the Wall in the PAC System: Displays 5 and 6
The approach obviously requires a thorough knowledge of the alphabet, and in practice users end up mumbling the alphabet under their breath, or asking the librarian for a copy of the alphabet. As a result, the instruction sheet shown in figure 8 is now displayed beside each terminal, with the alphabet printed vertically on the right. CLSI now provides keyboard terminals as an option as well.

**SEARCH COMMANDS**

In the command approach, the content of a search element is, of course, preceded by a command. But there is no uniformity or consensus on what a command language should look like. Here, for example, are the commands for an author search.

- **Find Author (or F A)**
- **F Au (or Auk)**
- **Find Personal (or Corporate)**
- **Find Pa (or Ca)**
- **A**
- **A =**
- **A/**
- **AU**
- **>AU**
- **AUT/**
- **PPNK**
- **author$**
Public Access Catalog Instructions

1. Touch **START OVER** to begin or end your search.

2. Touch the name of the file you wish to search: TITLE, AUTHOR, SUBJECT.

3. Touch the entry which comes alphabetically or numerically before the one for which you are searching. **Numbers file before letters.**

4. Repeat Step 3 until you see **BROWSE-REVERSE** and **BROWSE-FORWARD** at the bottom of the screen.

5. The entry you are looking for should now be listed on the screen. If not, the library does not have it, OR it is listed under another spelling or heading.

6. Now touch the term you selected to get a list of materials the library owns.

7. If there is more than one entry, continue to touch **SCROLL UP** to see them one by one.

8. When you see a title you want, touch it to get complete information, including whether copies may be on the shelf or checked out. (You must wait several seconds for this information.)

9. The screen now shows you the information the card catalog gave you. In addition you will see:
   - The code for main library or branch where item is located (code explained on reverse side of sheet).
   - Item number for each copy owned.
   - Indicates that this copy is checked out and is due back on the date shown.
   - Indicates this copy may be on the shelf. Write down the call number for locating it. Call number is in the upper left hand corner of screen.
   - If the call number says **ONORD** (on order) or **REC'D** (received), the material has been ordered or has been received for the Library's collection.
   - Touch **PAGE NEXT** to see if there is more information about this item. (Repeated beeps mean all information is on the first screen.)

10. Touch **START OVER** to end your search or to begin another. If you need assistance, ask a librarian or touch **HELP**.

Note that in the first four instances, the command "find" is explicit, either in full form or in abbreviated form (in some cases, it may be spelled out or abbreviated). In another six cases, the command "find" is implicit, followed by an abbreviation for the search element (in this case, "a" or "au" or "aut"), sometimes with an accompanying symbol after or before the search element. In another case, there is simply a group of letters, without much mnemonic value. And in the final instance, since
the word *author* does not appear in the command, there are *only* symbols, which are placed before and after the content of the search element.

In title searching, at least six systems treat the words as keywords, and retrieve all titles containing those words anywhere in the title. This method can be very useful, but it can also retrieve more titles than wanted. In the California system, a search for *War and Peace* retrieves 495 titles! Three systems—LC, Mankato, and Washington State—have a title command that involves a key constructed from the first four words, which effectively specifies that the words be in that order. In the LC system, for example, a search for *The Sound and the Fury* is made thus:

```
PTK SOU, A,T,F
```

Most systems drop the initial article automatically, but DataPhase, Northwestern, Ohio State, and Universal require the user to do this; DataPhase, in fact, instructs the user to try it without the initial article, and if the title is not found, to try it with the initial article.

For subject searching, at least six systems require that the exact LC subject heading be used, with the commas and dashes in place. Northwestern and Ohio State encourage the user to search the printed LC subject heading list first.

Several systems allow "combination" searches. The most common is an author/title key, composed of three to ten characters of the author's name, and four to ten characters of the title. Claremont has a special command that allows the user to combine an author field and/or a title with a subject term, in a predetermined format. Figure 9 shows a search using this command. In the menu option of the California system, the subject search actually searches for the words entered in both subject headings and in titles, so that the user normally finds something even if the term used is not in the LC subject heading list. Similarly, the author search also searches automatically for the author's name used as a subject as well.

Four systems allow the user to search for a term in any indexed field; in the Dartmouth system, in fact, the fields that may be searched include practically the full MARC record, except for a few fixed fields. The user can also search the notes in the Beth Israel and LC systems; by the publisher in Washington State; and by the name of the course and/or instructor (for reserve books) in Geac and ULISYS.

**Limiting and Broadening Searches**

In many systems, the "find" function is only one of several searching techniques. Of the other techniques, the most useful and therefore the most common are those that allow the user to limit a search or broaden it. There are at least five such functions: Boolean logic connections; a "backup" function; limiting by another search element; limiting by library; and truncation.

The use of Boolean connectors—that is, joining search terms with a logical "and," "or," or "not"—either limits the search or broadens it, depending on which connecting word is used ("and" and "not" limit
the search, "or" broadens it). In the following search, the use of "and" narrows the search to a manageable number of records.

**FIND SU SAILING**

Search request: **FIND SU SAILING**
Search result: 99 records at UC libraries

**AND SU OCEAN**

Search request: **FIND SU SAILING AND SU OCEAN**
Search result: 5 records at UC libraries

In most systems using Boolean logic, the combination need not be made in a stepwise fashion but can be made immediately, i.e., as a part of a single command.

Find su computers and su libraries

**SSOLAR ENERGY + RESEARCH**

find s water conservation not s soil conservation
find s banking and (s maryland or s pennsylvania)

As indicated in table 3, half of the twenty systems provide some Boolean logic capabilities. Note that this technique is much like the preformatted "combination" searches discussed earlier, except that the user may choose the particular combination of indexes and terms to be made.
### TABLE 3
**METHODS OF BROADENING AND LIMITING SEARCHES IN TWENTY ONLINE PUBLIC CATALOGS**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Avatar Systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beth Israel Hospital</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biblio-Techniques</td>
<td>x†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of California</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carlyle Systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claremont Colleges</td>
<td>x†</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLSI</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dallas Public Library‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dartmouth College</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DataPhase Systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geac International</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inforonics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library of Congress</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mankato State University</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Northwestern University</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio State University</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pikes Peak Library District‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal Library Systems‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington State University</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington University School of Medicine</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Boolean = "and," "or," or "not" command. Date = publication date or range of dates. Lang. = language of text. Lib. = holding library. Backup = command to return to previous state of search. Trunc. = omission of final letters in command. "Wild Cd." = omission of random middle letters.

† Searches by predetermined combinations of search elements are possible.
‡ These systems do not use these methods of broadening and limiting searches.

---

A "backup" command expands the search. This is particularly useful when a Boolean command has limited the search too much; the "backup" command then allows the user to return to the previous state of the search without having to begin the search again, as illustrated here.

**find su energy**

Search request: **FIND SU ENERGY**

Search result: 2710 records at UC libraries

and su environment

Search request: **FIND SU ENERGY AND SU ENVIRONMENT**

Search result: 4 records at UC libraries

Backup

Previous results are:

Search request: **FIND SU ENERGY**

Search result: 2710 records at UC libraries

Limiting by adding another search element, such as date and language, is very much like a Boolean "and," except that in most systems it is a separate command. Seven systems allow the user to limit by date, and
five by language (see table 3). In addition, the online union catalogs allow the search to be limited to a particular library, a technique that can also help to narrow down a search.

Truncation is a means of expanding the search; a string of characters retrieves any term starting with those characters. Sixteen systems provide this capability. In seven, the truncation is implicit; that is, all terms beginning with the characters entered are automatically retrieved whether this result is wanted or not, and a selection must then be made from among the terms retrieved. In the other systems, the truncation must be made explicit, usually by adding a pound sign after the string of characters. For example:

Search request: FIND SU LIBRAR# AUTOMATION
Search result: 136 records at UC libraries

Only one system—Mankato—provides so-called middle truncation. Middle truncation is really the use of a symbol as a “wild card” character; for example, “wom?n” retrieves both “woman” and “women.” This feature can help deal with spelling problems; if the user can never remember how to spell “receive,” for example, he can type “rec??ve.”

Another aid to searching is the so-called browsing function. In five systems, a command such as “browse” allows the user to see the indexes themselves, sometimes with cross-references. Here is an example.

Browse request: BROWSE SU VOLCAN#
Browse result: 140 subject headings found in the subject index

1. Birds — Lassen Volcanic National Park
2. Camping — Hawaii — Hawaii Volcanoes National Park — Guide-books
3. Hawaii Volcanoes National Park — Guide-books
4. Hekla (Volcano)
5. Hiking — Hawaii — Hawaii Volcanoes National Park — Guide-books
6. Lassen Volcanic National Park
7. Lassen Volcanic National Park — History
8. Mud volcanoes — Azerbaijan — Lok-Batan
9. Myojin-sho (Volcano)
10. Nyamplagira (Volcano)
11. Nyiragongo Volcano
12. Paricutin (Volcano)
13. Seismological research — Hawaii Kilauea Volcano
14. Seismology — Hawaii Kilauea Volcano
15. Soufriere, Guadeloupe (Volcano) — Eruption, 1798
16. Teneuogia (Volcano)
17. Tetela de Volcan, Mexico — Economic conditions
18. Tetela de Volcan, Mexico — Politics and government

Press RETURN (or type NS) to see the next screen.

In seven other systems, this function is implicit in the way the search operates: the result of a search is always a display of the term surrounded by adjacent terms in the index, from which a specific term may be selected.
A unique searching aid is the "keyword approximation search" at the Washington University School of Medicine. If nothing is found under the term entered, the system drops trailing s's, double consonants, and vowels, then attempts to match the remaining consonants with terms in the indexes. In effect, it operates as a spelling-checker. In the following example, surgery is misspelled, but the system nevertheless completes the search.

BOOKS BY KEYWORD / DATE:
ENTER KEYWORD(S), RECORD # OR RETURN FOR EACH OF THE FOLLOWING FIVE ITEMS:
(DO NOT USE ABBREVIATIONS OR INITIALS—USE ONLY WORDS OF WHICH YOU ARE SURE)

<table>
<thead>
<tr>
<th>AUTHOR</th>
<th>PARKER</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE</td>
<td>SERGERY</td>
</tr>
<tr>
<td>SUBJECT</td>
<td>(ANY)</td>
</tr>
<tr>
<td>SERIES</td>
<td>(ANY)</td>
</tr>
<tr>
<td>DATE(S)</td>
<td>1900 -</td>
</tr>
</tbody>
</table>

TRYING KEYWORD SEARCH: — NO BOOK CONTAINED ALL KEYWORDS
TRYING KEYWORD APPROXIMATION SEARCH:

<table>
<thead>
<tr>
<th>RECORD #</th>
<th>CALL #</th>
</tr>
</thead>
<tbody>
<tr>
<td>35472</td>
<td>WE 17 W564a 1981</td>
</tr>
<tr>
<td></td>
<td>Wheeless, Clifford R.</td>
</tr>
<tr>
<td>359L9</td>
<td>WO 11 FA1 P239e 1920</td>
</tr>
<tr>
<td></td>
<td>Parker, George</td>
</tr>
</tbody>
</table>

Record Displays

Once a search has been completed through one or more of these techniques, what happens? In many of the systems, there is an automatic display in a particular format. In eight of the systems, however, the results are reported first, so that if more records have been retrieved than are really wanted, one of the limiting devices mentioned earlier can be used. Conversely, if too few have been retrieved, the search can be broadened. In six of these eight systems, the search is also confirmed, so that the user knows how the system interpreted what was typed.

find su mistaks
Search request: FIND SU MISTAKS
Search result: 0 records at UC libraries

The number of different display formats used in a single system varies from one to five. As indicated in table 4, two systems have only one; seven systems have two; eight systems have three; two systems have four formats each; and one system, Mankato, has five, including two different forms of what I have called "brief."

An "index" display, usually consisting of only one or two lines, is found in fourteen systems. Two examples are shown in figure 10. Some of them do not include the call number, and a separate screen must be requested to obtain that information.
TABLE 4
TYPES OF DISPLAY IN TWENTY ONLINE PUBLIC CATALOGS

<table>
<thead>
<tr>
<th>Online Catalog</th>
<th>Index</th>
<th>Brief</th>
<th>Full</th>
<th>MARC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avatar Systems</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Beth Israel Hospital</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Biblio-Techniques</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of California</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Carlyle Systems</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claremont Colleges</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dallas Public Library</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dartmouth College</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>DataPhase Systems</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geac International</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inforonics</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library of Congress</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mankato State University</td>
<td></td>
<td>x</td>
<td><strong>x</strong></td>
<td>x</td>
</tr>
<tr>
<td>Northwestern University</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio State University</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pikes Peak Library District</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal Library Systems</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington State University</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington University School of Medicine</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Two different forms of brief entry are available.

Sixteen systems have a brief bibliographic display, usually including brief author, title, imprint, and call numbers.

LO DRBU.

GF75-C65-1974.

ME Commoner, Barry, 1917-

TI The closing circle: nature, man, & technology / Barry Commoner.

IM New York: Bantam Books c1974

SU Pollution

Search Request: FIND PA COMMONER AND TI CLOSING CIRCLE

Search result: 2 records at UC libraries

1. Commoner, Barry, 1917-
   The closing circle; nature, man and technology. 1st ed.
   New York, Knopf, 1971
   UCI Main Lib GF75.C65
   UCSC McHenry GF75.C65
   UCSD Central GF75.C65
   UCSD Cluster GF75.C65

2. Commoner, Barry, 1917-
   The closing circle; nature, man, and technology.
   New York, Bantam Books 1972, c1971
   UCSD Central GF75.C65 1972

Eighteen systems have a full bibliographic display, with all or most of the bibliographic information found on a catalog card. In only one or two cases, however, is the information arranged as it is on a card. In most systems, the information is in tabular form, with the different parts labeled.
Characteristics of Online Public Catalogs

Table of Contents


xvi, 200 p. : ill. ; 24 cm.

Series: New directions in librarianship ; no. 2. 0147-1090
Notes: Includes index.
Subjects: Machine-readable bibliographic data — Congresses.
Library catalogs — Congresses.
Libraries — Automation — Congresses.
Library administration — Congresses.

Other entries: Spyers-Duran, Peter.
Gore, Daniel.
Kimbrough, Joseph.
Associated Colleges of the Midwest.

Call Numbers:
UCB LibSchLib Z699.A1R46 (CU-LSL)
UCD Main Lib Z699.A1 A18 1977 (CU-A)
UCI Main Lib Z699.A1 A18 1977 (CU-I)
UCLA Eng/Math Z699.A1 A18 1977 (CLU-EMS)
UCLA URL Z699.A1 A12 1977 (CLU-URL)
LBL Main Z699.A1 A2 1977 Bldg 50
11 Lib./Off. (CU-Lbi)
UCR General Z699.A1 R46 (CU-Riv)
UCSB Library Z699.A1 A18 1977 (CU-SB)
UCSD Central Z699.A1 A18 1977 (CU-SCu)

Some recent research seems to indicate that this kind of format is preferable to the narrative, or paragraph, form of the catalog card, and makes it possible for users to interpret the information much more quickly. Particularly to be avoided are displays like the following, which is not only strung together in a run-on fashion, but interlarded with extraneous and distracting symbols.


Five systems also provide a full MARC display, in some cases showing the variations in cataloging among member libraries as well. Figure 11 is an example from the California system, which takes six screens. Presumably only library staff are interested in such displays.

Two systems display the authority file, and one of them, Bibliotechniques, actually has three authority file displays: one with the headings only, one with the cross-references and notes, and one complete with tags.

Finally, two systems, California and Bibliotechniques, allow the user to display certain selected fields alone. Here, for example, is a display from the California system showing subject fields only.

Search request: FIND SU LIBRAR# AUTOMATION
Search result: 136 records at UC libraries
Library catalogs — Congresses.
Libraries — Automation — Congresses.
Library administration — Congresses.

2. Subjects: Library catalogs on microfilm.
Computer output microfilm devices — Library applications.

Libraries — Mechanical aids.
Libraries — Automation.

Press RETURN (or type NS) to see the next screen.
Characteristics of Online Public Catalogs /57

Search request: FIND SU LIBRARY AUTOMATION
Search result: 136 records at UC libraries

1. 000017068 STA p REC an ENC DCP i CSC EMT 780830
   INT 1 001 GOV PSC 0 BIO CIV etc COM ILS a
   REF MEL 1 CMF 1 FTC 0 MOG LAN eng PO sl979
   020 313206082 : $c $17.50
   035 $# 0001 $a INV/CRGC $c 04496165 $s 810210 DAV-G
   035 $# 0002 $a INV $c 04496165 $s 0137568 $s 810210 IRV-G
   035 $# 0003 $a SB $c 04496165 $s 810210 SB-G
   035 $# 0004 $a LA/CLAIRE $c 04496165 $s 810210 LA-G
   035 $# 0005 $a CLAIRE $c 04496165 $s 810210 LA-P
   035 $# 0006 $a SD/CUSL $c 04496165 $s 810210 SD-G
   035 $# 0007 $a BER $c RL1354266 $s 61494793 $s 810410 BER-G
   035 $# 0008 $a DAV $c RL14357480 $s 5160780 $s 810410 DAV-G
   035 $# 0009 $a LBL $c RL19883099 $s 810410 LBL
   040 DLC $c DLC $d DAIV-G
   040 DLC $c DLC $d CU-RIV
   040 DLC $c DLC $d UC-BER
   040 DLC $c DLC $d UC-LBL
   (Record 1 continues on the next screen.)
   Press RETURN (or type NS) to see the next screen.

1. (continued)

   040 DLC $c DLC $d CU-SB SB-G
   040 DLC $c DLC $d CLU-LA,LA-G
   040 DLC $c DLC $d CU-LA
   040 DLC $c DLC $d UC-BER-LB
   040 DLC $c DLC $d UC-LBL
   050 0 E699.AI $b R46 BER-G,RIV-G
   090 $# 0007 $c 04495416 5 $e BER $c UC-COL $q 51649793 $z 810410 BER-G
   090 $# 0008 $c 044954165 $e DAV $c UC-LA $q 5160780 $z 810410 DAV-G
   090 $# 0009 $c 044954165 $e LBL $c UC-LA $q 5160780 $z 810410 LBL
   100 (Record 1 continues on the next screen.)
   Press RETURN (or type NS) for the next screen. Type PS for the previous screen.

1. (continued)

   090 $# 0002 $c 044954165 $b A18 1977 $c INV $c UC-I $q Main Lib $b 51649793 $z 810410 INV-G
   090 $# 0003 $c 044954165 $b A18 1977 $c LA $c CLU-MATH $b 51649793 $z 810410 LA-P
   090 $# 0004 $c 044954165 $b A12 1977 $c LA $c CLU-URL $b 51649793 $z 810410 LA-G
   090 $# 0005 $c 044954165 $b A2 1977 $c UCL $c UC-LA $q Main $b 51649793 $z 810410 UC-LB
   (Record 1 continues on the next screen.)
   Press RETURN (or type NS) for the next screen. Type PS for the previous screen.

1. (continued)

   090 $# 0006 $c 044954165 $b A18 1977 $c SE $c CLU-CSC $q Central $b 51649793 $z 810410 SE-G
   111 0 A00 Conference on Management Issues in Automated Cataloging, $c Chicago, $d 1977.
   245 10 Requiem for the card catalog : $b management issues in automated cataloging / $c edited by Daniel Gore, Joseph Kibrough, and Peter Spyers-Duran. BER-G, RIV-G
   300 $x 200 p. : $b ill. ; $c 24 cm.
   (Record 1 continues on the next screen.)
   Press RETURN (or type NS) for the next screen. Type PS for the previous screen.

1. (continued)

   300 $v 200 p. : $b ill. ; $c 24 cm. BER-G
   400 0 New directions in librarianship ; $v no. 2. $x 0147-1090
   440 0 New directions in librarianship ; $v no. 2 $x 0147-1090 LA-P
   440 0 New directions in librarianship ; $v no. 2 $x 0147-1090 LA-P
   500 Papers of a conference held in 1977 and sponsored by the Associated Colleges of the Midwest. BER-G, RIV-G
   504 Includes bibliographical references and index. LA-P, RIV-G
   Press RETURN (or type NS) for the next screen. Type PS for the previous screen.

Figure 11
MELVYL's Full MARC Display
Figure 11 (Continued)
MELVYL’s Full MARC Display

This feature can be particularly helpful if the user is trying to determine what the proper LC heading is. For example, the popular term can be used in a title search; then, if some records are retrieved, the subject fields can be displayed, and the appropriate subject headings noted; and a subject search can then be made on these terms. The feature is also useful in discovering related terms that might not have occurred to the user.

USER ASSISTANCE

So far this paper has discussed the way these systems get the user started, allow searching of various bibliographic fields, and display the results. In most cases, however, these functions are not enough to ensure that the user can find what is wanted easily and effectively. In addition to the various searching and display techniques, therefore, most of the catalogs incorporate additional methods of what may be called, generically, user assistance.

Probably the most helpful means of assistance is the use of common English, and the avoidance of jargon, computerese, symbols of the type commonly used by programmers, and what Neville has called “Bibliish”—that is, the kind of cant perpetuated by catalog cards and other bibliographic records. A few examples follow; keep in mind that these instructions are ostensibly for users who encounter the catalog and computer terminals relatively infrequently.

THE ONLINE CATALOG IS A MODULE WITHIN THE CIRCULATION SUBSYSTEM, AND IS INTENDED FOR PUBLIC USE. YOU MAY USE A LIGHT PEN TO READ THE BARCODE NUMBER OR TYPE IN THE BARCODE NUMBER AND PRESS CONTROL AND "R" SIMULTANEOUSLY.

FOR “FAST SCANNING” RECORDS COLLECTED WITH THE “G,A,T,Y, OR S” CODES, INSERT “F” (FAST FORWARD) OR “R” (FAST REVERSE) IN 2ND CHARACTER POSITION.
The optional field-type identifiers can be used following a logical connector and preceding the word or words they pertain to.

The format requested was unclear. The command should be entered:

```
Format:   D SET#/FORMAT/DISPLAY OPTION
Example:  D 1/L/TTL/ABST
```

YOU HAVE ATTEMPTED TO CHANGE THE TYPRUN OPTION, WHOSE VALUE IS FIXED FOR YOUR TERMINAL AND CANNOT BE CHANGED. THE VALUE OF THE TYPRUN OPTION FOR THIS TERMINAL IS "$TYPRUN."

INITIALIZATION COMPLETE

Another difficulty with language is that it can be either too terse or too verbose. If it is too terse, the user is left wondering at the meaning, and perhaps thirsting for more information. Imagine a user confronted only with the following instructions:

Look for:
Enter inquiry request?

If it is too verbose, on the other hand, the user rapidly grows impatient at having to plow through all the text before getting on with it.

Here are some examples of language that is a bit more friendly.

To search the ILS online catalog you must first identify the type of search you wish to perform. For instance, if you want items by a given author, choose the AUTHOR search; if you know the title of the desired item, . . .

To search for books in the catalog, type FIND, followed by the index you wish to search, followed by the keywords you want to search in that index.

Your search, FIND SUBJECT ENERGY, found 136 records at Colleges of Mid-America.

You may choose the type of search you want by touching the appropriate word on the screen.

To find out what to do now, type OPTIONS (or O)

I don't have an explanation for that topic. Did you type it wrong, perhaps?

Ask your librarian for assistance.

In addition to making the language clear, simple, and not too chatty, there are at least five other methods of assistance that an online catalog can provide: editing, prompting, error messages, help messages, and tutorials.

Provision for simple editing is one of the most common methods of assistance. Usually the system is programmed so that the backspace can be used to erase mistakes; at least seven systems do this. Additionally, some systems allow the user to push a "break" key to stop a display and go on to whatever comes next, clearly a helpful feature for people who are impatient or in a hurry.

Another kind of assistance is the prompt message, examples of which follow.

ENTER AUTHOR'S LAST NAME
DO YOU WISH TO ADD TO YOUR SEARCH? (YES/NO)
FOR A REVIEW OF COMMANDS TYPE "HELP" OR "?"
Type SUBJECT words below or type HELP, then press RETURN
Press RETURN to see the next screen

Error messages are comments made by the system when the user makes an error in format, or in spelling a command, or in understanding the logic of the system. As in most other respects, the systems now available vary widely in the kind and number of error messages supplied. Surprisingly, five or six of the systems provide no error messages. If the user makes a mistake, nothing happens, or the system simply repeats the previous instructions until the user gets it right. In four or five others, the system simply says "invalid command," which is not particularly helpful. One step above this are the systems which provide at least a few error messages that begin to suggest what went wrong; typically, these systems also say "invalid command," but may also say things like this:

No entries found
No match found—try again?
No copies exist in this agency

A step above this are the systems which provide a variety of error messages tailored to the nature of the error, but unfortunately in bibliographic or computer jargon that may not be very effective in helping the user correct the mistake.

****CCMB0002— SCORPIO COULD NOT FIND AN EXPECTED OPERATOR OR RIGHT PARENTHESIS AFTER ONE OF THE SET NUMBERS IN THE LAST COMMAND. CHECK FOR BALANCED PARENTHESSES, MISSING OR MISSPELLED OPERATOR.
OPERATOR OR RIGHT PARENTHESIS?
FORMAT: COMB SET# OPERATOR SET#
EXAMPLE: COMB 1 AND 2
READY FOR NEW COMMAND:

The human engineering studies mentioned earlier suggest that error messages should do at least two things: they should indicate specifically and clearly what went wrong; and they should indicate what the user can do about it. Of course they should do so in clear English as well. In addition, they should be positive rather than condemnatory, and attempt to reinforce the notion that the user is in control. Here are a few examples that come closer to that goal.

A modifying command (such as AND, AND NOT, or OR) cannot be used to start a search. Use FIND instead.
Your command, FIND SUBJECT, does not include any terms to search. Please retype.
Your command must include an index name after FIND.
Type HELP for valid index names.
You must perform a BROWSE before you can SELECT any fields.
To look at your search result, use the DISPLAY command.

Related to the error message is the help message. Whereas error messages appear automatically when an error is made, help messages nor-
nally appear only if the user asks for them, usually by typing "help" or a question mark, or selecting the "help" option from a menu.

At least three systems have no help messages. Three or four others have only one, usually an introductory screen explaining the commands recognized by the system, and perhaps one other general set of instructions. Here is an example.

**PPL HELP**

Command choices:
- A/data = Author search
- B/data = Barcode search
- C/data = Call number search
- E/ = Expand this record
- F/ = Forward
- L/data = LCCN search
- N/ = Next author/title
- S/data = Subject search
- T/data = Title search
- X/ = Exit program

Eleven of the twenty systems have a more extensive inventory of help messages, ranging from seven in the Northwestern system to a high of 155 in the California system. In the more elaborate systems, help messages fall into two categories: contextual help messages, which appear after the command "help" and provide information specific to the point in the search session where you’ve asked for help; and addressable help messages, which provide information on specific topics and can be called up on command from any point in a search.

Contextual help messages normally occur when an error message has been received, and the user then asks for more information by typing "help."

---

**FI SU WAR AND PEACE**

SEARCH REQUEST: FI SU WAR AND PEACE
SEARCH RESULT: 159 records at UC libraries

**DIS FULL**

The word FULL cannot be used in a DISPLAY command

**HELP**

You have typed a DISPLAY command which contains a word that is not recognized by the system. Please retype your command, making sure that you have used the correct command words.

For further instruction, type:

**HELP DISPLAY**

Find su warr
Search request: FIND SU WARR
Search result: 0 records at UC libraries

**HELP**

Your search, FIND SU WARR
does not retrieve any books from UC libraries. First check the spelling of your
search request to be sure that MELVYL is looking for what you really want. Then be sure you are searching the correct indexes.

Even if you do not retrieve any books, what you want may still be in the library. The on-line catalog contains records for only a small percentage of the books in the University. Check your local catalog for a more complete representation of the books in your library.

For further details in the construction of search requests, type:

```
HELP SEARCH HELP INDEX HELP KEYWORD
HELP CONJUNCTION HELP TRUNCATE
```

The following two examples are addressable help messages. In the first, the user has asked for help on commands, and in the second the user has asked for more information on the display formats available.

HELP COMMAND
A command is an instruction given to TOMUS. After typing a command, press the RETURN key to let TOMUS receive the command.

Every command must begin with a valid word. The following words are valid:

```
AND DISPLAY LOGOFF SELECT STOP
AND NOT END OR SET
AT FIND PS SHOW
BROWSE HELP QUIT START
```

Any command word may be abbreviated down to 3 letters. FIND, HELP and DISPLAY may be abbreviated to 1 letter—that is, F, H and D.

If you try to type a command that does not begin with one of the words listed above, TOMUS will send an error message to explain what went wrong.

For further information about any of these commands, type HELP followed by the particular command; for example, HELP FIND.

HELP DISPLAY
After your search has found some records, use the DISPLAY command to see the records. Type DISPLAY followed by the number or numbers of the records you want to see. If you type DISPLAY without any record numbers, you will see all of the records.

To see more than one record but not all of them, use a blank or a comma to separate the numbers; for example, DISPLAY 1 3 5, or DISPLAY 1,3,5. To see a range of numbers, use a hyphen; for example, DISPLAY 10-20.

You may also specify another display format. The normal format contains the basic bibliographic information on the book, the call number, and the location. The other formats are BRIEF, FULL and MARC.

- **BRIEF** consists of only one or two lines per record. Use it to look at a large number of records to decide which ones to look at more closely.
- **FULL** consists of all the information in the record, including notes.
- **MARC** contains the computer tags and other technical information.

For more information on the formats, type HELP followed by the format name.

The command DISPLAY may be abbreviated to DIS, DI, or simply D.

Some examples of valid DISPLAY commands are:

```
DISPLAY 1 3 5 FULL
DIS 1-10 D 3 6 MARC D FULL
```

In the California system, there is also a provision for automatically recognizing when a user may be in trouble, and providing help. If the same error is made three times in succession, the appropriate help message is automatically displayed.
find energy
Your command must include an index name after FIND.
Type HELP for valid index names.

find energy
Your command must include an index name after FIND.
Type HELP for valid names.

find energy
A FIND or BROWSE command must always include a primary index name after
the command word FIND or BROWSE. This tells MELVYL which index to
search.

You may use only the following two-letter abbreviations for the primary indices:

PA  (personal author)
CA  (corporate author)
TI  (title)
UT  (uniform title)
SE  (series)
SU  (subject)

For explanations of the contents of each of these indices, type HELP followed by
the two-letter index name.

A final method of user assistance is the provision of a tutorial mode, or
session, that takes the user through a predetermined set of instructions.
No system discussed in this survey has incorporated a tutorial mode.
The closest is Northwestern, which has nine introductory screens. A tu-
torial mode is under development at the University of California; how-
ever, experience with that system seems to indicate that the help screens,
which can be addressed as needed or wanted, form a kind of user-defined
tutorial, and may be sufficient. Charles Hildreth has suggested that
"learning while using the system, guided by suggestive prompts, con-
structive error messages, and . . . help displays, may well become the
preferred method of self-training."

**DATABASES, EQUIPMENT, AND COSTS**

There are some fascinating bibliographic questions involved in these
online catalog systems, but unfortunately space does not permit a full
discussion here. It should be pointed out, however, that only eight of the
systems have some form of authority control, at least at present, and
only three of these consolidate variant records for the same title, and re-
tain those variations in the master file. I have discussed some of the prob-
lems and techniques of authority control and consolidation in another
paper.

The bibliographic databases in these systems are created and main-
tained in three different ways: by loading and reading tapes from
OCLC, RLIN, or some other bibliographic utility; by an online inter-
face with a utility; or by keying the information directly into the system
(see Table 5). Sixteen systems load tapes to create the database, and four-
ten use the same method to maintain it. For six systems, an online in-
terface to OCLC is available. Thirteen permit direct keying of records
into the system and online editing; for the Dallas Public Library this is
the only method of input, and of course, LC creates all of its own records as well.

The size of the database varies greatly from system to system, although the sizes tend to cluster in certain ranges. There are two of approximately 35,000 records; seven clustered in the 400,000-500,000 range; and three between 700,000 and 750,000. The others are scattered between 69,000 and 2,500,000.

Nine systems use small processors, such as a Data General, a PDP-11, or an IBM Series 1; eleven use large machines, usually an IBM or an IBM plug-compatible machine, although one system (Mankato) uses a Univac 1100 series. Most will work with any ASCII terminal, although five use special terminals (see table 5).

As with many automated systems, costs are sometimes difficult to assess, or are even unknown. However, some generalizations can be made from partial information. Those that are available for sale cost several hundred thousand dollars: in four cases, from $200,000 to $500,000, and in three cases over half a million. Software only costs between $50,000 and $150,000, if you have your own equipment or want to buy it separately. One system, Carlyle, is available on a service basis, like a utility, at $400 per month per port (a port being a telecommunications link between the library and the computer center), and others can probably be negotiated on a lease or lease/purchase basis. Development costs,
where they are known or available, range from half a million to several million dollars.

CONCLUSIONS

Several general conclusions about existing online public catalogs can be made. First, of course, the number of online systems is growing. The field is also evolving rapidly, a fact best seen by comparing the early systems with later ones.

Second, neither of the two basic approaches to searching—menu and command—has become dominant. In fact, libraries and users seem to prefer having both, and some of the systems obligingly provide both.

Observation and comparison of the systems can also lead to some tentative conclusions about what characteristics online systems should have. Such conclusions are admittedly risky and subjective, and perhaps should be considered suggestions or notes toward a philosophy of online catalogs, rather than fixed conclusions based on evidence.

• Menus should probably be mnemonic where possible, rather than numeric.
• Commands should use common, descriptive English words, or mnemonic abbreviations, or both. They should also imply that the user is in charge, and be “friendly” in tone.
• The language of the system should be simple, plain, and free of jargon. It should also be relatively terse, but with branches to more extensive text if the user wants it.
• Instructions to the user should include examples.
• Initial articles of titles should be dropped automatically.
• Keyword searching is preferred, especially for subject searching.
• Boolean connectors, limiting terms, and truncation are helpful in broadening or limiting searches.
• Browsing the indexes is a great help, especially in subject searching.
• Display formats should probably include an “index” format, with one or two lines, and a “brief” format. A full bibliographic format is also desirable. Less clear, however, is whether a MARC format is really needed.
• User assistance devices should include prompts, error messages, and help messages.
• Prompts should be very brief, so as not to delay the experienced user unnecessarily.
• Error messages should indicate specifically what went wrong, and what can be done about it. They should also be positive in tone and free of jargon.
• Help messages should appear in answer to help requests at any point, and provide specific help suggestions appropriate to that point in the search.
• Hardware and software should be adequate to handle the size of the data files expected, and provide responses within about five seconds.

Those who are interested in pursuing idealism further may want to consult the very elaborate specifications for an online catalog prepared
by the Consortium to Develop an On-line Catalog, or CONDOC, with the assistance of Joe Matthews and Associates. These are available from the chair of the consortium, Robert Miller, at the University of Notre Dame Libraries.

The diversity of approaches used by these twenty systems can be bewildering, especially if one is trying to absorb the differences all at once. It may also give the impression that it's all too complex for the average user. It must be kept in mind, however, that the user is not presented with twenty different systems; the user has only one to contend with, and normally finds it very easy to understand. A study of online catalogs funded by the Council on Library Resources and carried out by researchers at OCLC, Research Libraries Group, the Library of Congress, the University of California Division of Library Automation, and Joe Matthews and Associates provides support for this conclusion. The study included some of the earlier and less friendly systems, but nevertheless found that 92 percent of the users like the online catalog. Almost 75 percent think it’s better than the card catalog. These percentage figures are averages, of course, and for some of the better-designed catalogs the numbers are even higher. The comments are even more encouraging: users say things like “it’s fantastic!” “the greatest thing since sliced bread,” “I found 135 books on my subject in less than two minutes,” “cuts research time in half,” “terrific,” and perhaps the most common comment, “it’s fun!”

Users like the online catalog not only because it’s fun, but also because it’s interactive—instead of just sitting there, it talks back, makes suggestions, and in some of the better systems seems to be genuinely interested in helping the users find what they want. It gives the users a sense of control, and also a sense of accomplishment, because it provides so many ways to reach the goal; many, if not most, of the functions described could never be performed with a card catalog.

Any librarian who has watched the public using an online catalog cannot help but be excited at the response. Librarians have been trying for centuries to help users find what they want, and online catalogs provide a major step forward in doing this. Our challenge now is to use this tool, and refine it, so that our services rise to a new and higher level.

REFERENCES


6. University of California, Division of Library Automation and Research and Analysis Group, "Users Look at Online Catalogs: Results of a National Survey of Users and Non-Users of Online Public Access Catalogs; Final Report to the Council on Library Resources" (Berkeley, Calif.: Univ. of California, Systemwide Administration, 1982), p.92, 95.
The Human Dimension of the Catalog: Concepts and Constraints in Information Seeking

Leigh Estabrook

This paper concerns the social and psychological effects on users of the introduction of the online catalog in libraries. Within that framework, the argument put forth is that users' access to information is inextricably bound up with the social context within which it is made available. Conversely, our decisions about how to provide access to information are bound up with our social and cultural systems.

Let me begin, however, with two preliminary points. First, I would argue, technology is not inevitability. Although libraries have become heavily committed to the online catalog as a solution to certain of their technical and economic problems, the adoption of any specific technology is not predetermined. Indeed, it is not a given that the solution to a library's problems is technological at all. Take, for example, the problem of bibliographic access for public users—basically, the identification of materials that are needed for and useful to research, school assignments, or general information problems. The online catalog is one solution to that problem and the profession continues to seek ways to improve its utility through enhancement of Library of Congress subject headings, expansion of bibliographic points of access, and other means. Alternatives to this technologically-intensive solution do exist. It could be possible, for example, for a library to hire a cadre of unemployed or underemployed Ph.D's, in a variety of subject areas, whose responsibility it would be to read reviews and research, to peruse bookstores, to attend conferences, and to engage in other activities that would enable them to act as specialists with extensive subject knowledge to be tapped by users. (I don’t want to get into a debate about the relative merits of this labor-intensive solution versus the computer: my purpose is only to stress the need to recognize that computerized systems are not the only possible solution to organizational problems.)

Leigh Estabrook, associate professor, School of Information Studies, Syracuse University, presented this paper on July 9, 1982, at the “Prospects for the Online Catalog” preconference.
My second preliminary point is that users come in two types. When we talk of users of online catalogs, we cannot limit ourselves to public users. It is necessary to talk about professional (or organizational) users also. For there are both intended and unintended consequences on both groups of the implementation of an online catalog. Both public and professional users are users of and used by the technology, and as a consequence relationships among and between public and professional users are affected.

In looking then at the social and psychological effects of the introduction of the online catalog, I assume that we cannot separate our technological choices from the social matrix in which those choices are made and the social contexts of their implementation. These points may be made most clearly if we examine several of the major changes to public and professional use of the catalog that result from online systems and in turn examine some of the unintended social effects of each of these changes.

**Effects of the Online Catalog on Public Users**

For public users, there seem to be three major differences between the online catalog and hard-copy/paper or microform formats: first, most obviously, online catalogs are computerized; second, they are or have the potential to be distributed; and third, they may be unified.

**Computerization**

There is no need to elaborate the benefits of computerized access: increased forms of access to bibliographic records, opportunities to manipulate bibliographic access points, etc. Granted, kinks in many systems remain, but as they are worked out it is impossible to deny that online catalogs can provide access to an increased number of materials.

The psychological and social consequences of working with computerized systems relate both to the ways in which public users come to think about the information that they seek and to their expectations about what information the system is providing. Although we do not have any supportive data at this time, it would seem that lack of transparency of computerized systems may affect the way in which individuals think about the information with which they are presented. The way in which items are entered into the online catalog, the different ways in which they might be retrieved, rules for assigning subject headings, and the decisions of which even many professional librarians are not quite certain. And yet those decisions are crucial to the user's ability to retrieve needed information. This is not to say that some of those decisions are not also important in the design and use of a hard-copy catalog; but as online systems offer increased points of access, the need to understand the system becomes more crucial to its successful use.

The lack of transparency of the system interacts with users' tendencies to believe blindly in the results that it gives. A recent study offers some insight into this problem. Researchers, it is reported, took a sample of
well-educated people, each of whom scored very well in a test to estimate numbers (e.g., 791 x 843) without the use of paper and pencil or hand calculator. These individuals were then asked to make a series of estimates of answers to problems and to check their estimates against the actual results from hand calculators. The calculators were programmed to give answers increasingly at variance from the true answer. As we might expect, a high percentage of the subjects in the study did not question the calculators’ answers. Instead, they questioned their own computational skills. There is some hope—one seventh-grader questioned the results of the calculator on the first problem. When the researcher asked, “You think there might be some problem with the calculator?” the kid answered, “Of course, my calculator at home is always going out.” Perhaps the development of computer literacy in the next generation will result in more such iconoclasts.

When we combine this lack of transparency of the online system and the tendency for users of computers to trust excessively in their results, it would seem likely that users of online catalogs may in fact obtain less information than they seek and to believe in the completeness of their search even more than they might were they using paper catalogs.

I might also add a cautionary note about the current emphasis on designing “user-friendly” catalog systems. Once again, there is a trade-off. While it is important to develop systems that are forgiving and helpful to the user (how many of us have kicked some computer while we tried to tell it “off,” and it kept responding to our various symbols with a question mark?), the more user-friendly a system is, the more likely we are to endow it with powers it does not possess.

Distributed Systems

The second major difference between the online and hard-copy catalog is the potential for the online catalog to be available to public users at numerous remote locations. Although a book or fiche catalog can also be made available in multiple copies, there are many more limitations; a user with a terminal and modem, for example, can use the Syracuse Sulirs from his or her home.

Distributed systems provide many additional opportunities for users to obtain physical access to the catalog. At the same time, implementation of such systems has the potential for radically changing the ways in which libraries are used by the public and the relationships between public and professional librarians.

In their most fully developed form, distributed catalogs offer the user the opportunity to search for and request library materials without ever entering the library building. And I understand that at least one university library has planned a system for delivering materials to fulfill these requests. To many users, such a possibility is nirvana—one never needs to drive, walk, or otherwise become inconvenienced to go to the library to go through the rigmarole of locating and checking out materials. The trade-off for such convenience is that public users miss the new-books shelf, lose the opportunity to pick up a recent magazine or newspaper, and—perhaps most significantly—lose the chance to browse among re-
lated items. In a sense, users return to a completely closed stacks system.

It is important that we recognize these trade-offs because the social
effects are not trivial. Individuals who do not come to the library are less
likely to understand the nature and scope of work that goes on or to de-
velop loyalty and ties to the library. They also decrease their chances for
serendipitous encounters with literature, information, and the cultural
holdings of libraries. Users who do not browse may be led to think of the
world of knowledge one-dimensionally or linearly rather than relation-
ally. If I go to the shelf to search for materials on human/computer
interaction, for example, it is easy for me to scan the shelves above, be-
low, and behind me to get a sense of related materials (however much I
am tied to the linear classification system). The computer may allow me
to play around with call numbers and at least one browsing online sys-
tem exists (for a small specialized collection) that allows me to look at
related materials. I may use a thesaurus or list of subject headings to get
an idea of how a collection is organized. Any computerized system still
seems to force the user into linear approaches to information seeking.

I mentioned also the user/professional relationships in distributed sys-
tems. Individuals who use library services away from the library call
upon professional services only indirectly—through the professional de-
sign of the online catalog and development of delivery services. Knowl-
edge the professionals gain about these clients may be obtained through
analyses of online use. In each case, however, the machine acts as an
intermediary between human actions. Relationships between profes-
sional and public user become indirect, less personal, and more systema-
tized.

**Unified Systems**

Online catalogs also tend to integrate an increasing number of func-
tions from circulation to acquisitions. As a consequence, the public user
can learn not only whether a library holds a particular work, but also
whether the book is in the library and where it is located in the stacks. To
have this type of information is a clear advantage to many users. At the
same time, online catalogs that are used to perform a variety of library
functions provide fertile territory for disputes among different groups of
users.

The design and implementation of any type of computerized technol-
ogy raises political issues within the organization in which it is installed.
While theoretically computers may achieve the power to perform all the
operations that each user group might desire, in practice the expense
and the issue of access to certain types of information limit both the form
and content of information stored in and accessible through an online
system. User groups—different types of public users and different types
of professional users—compete with one another (although perhaps not
openly) for the inclusion of features they prefer in a system. The more
the online catalog is expected to perform functions of inventory control,
acquisitions, cataloging, and public access to the collection, the less
likely it is that public users will have a system that is best designed for
that public use.
EFFECTS OF THE ONLINE CATALOG ON PROFESSIONAL USERS

Computerization, distribution, and unification of online catalog systems also have psychological and social effects on those professional users in libraries.

Computerization

For a library worker, the computerized catalog tied to OCLC or to another utility saves duplication of effort by providing access to the cataloging records of other organizations. It also provides indirect access to some of the holdings of other libraries and the possibility of obtaining cataloging information more rapidly. The benefits are great; so, too, the costs.

Psychologically, those who work with the computerized catalog are distanced from the substance of their work. There is that lack of transparency of the computerized system—the problem of knowing what goes on behind the screen—of knowing what really is in the database. There is the loss of that intimacy with materials in the collection that comes from working closely enough with an item to be able to say what it is about for cataloging and classification.

Within libraries the introduction of the computerized catalog is gradually changing the shape of the technical services labor force. Although few libraries have fired professional catalogers to replace them with non-professionals, there is evidence that the majority of cataloging staff are increasingly hired from the nonprofessional labor force. This shift has resulted from the increasing routinization of work as less original cataloging is done within many libraries—a fact that should be noted by those concerned about librarians' resistance to change in the face of new technology. If we look at hiring patterns, however, and the changes in the work structure, we may want to consider the extent to which resistance to computerized and subsequently online cataloging may be a survival mechanism for those in the affected labor market.

Not only is the structure of the cataloging labor force changing but also the social context of the work. Workers are more subject to the pacing of the computer than to their own rhythms and interests. Although I have not seen much evidence yet of productivity goals for library staff, connection to the online catalog also makes it easier for worker performance to be measured. Finally, cataloging through an online system—sitting individually at a video-display terminal and talking to the big computer in Ohio or wherever—is an isolating experience, quite different from working in a group with several catalogers and sharing informally on working problems.

Unified Systems

Unified systems that tie together acquisitions, cataloging, and other functions also greatly improve access for professional users of the online catalog. I remember well (and with pain) one of my first jobs in an acquisitions department. It involved the traditional searching and verifying through catalogs, the LC-slip file, the on-order file, the in-process file,
the dead file, and I can’t remember what else. It was excruciatingly boring and error prone.

As indicated in my discussion of public users, however, implementation of unified systems raises important political issues of power among user groups (in this case, the different departments in the library) concerned with the design of the computerized system. Since not all desired features can be incorporated into any one system, conflicts can emerge over systems design. As a result, it may happen that computer programmers and software developers may have greater authority in the design of the system than might have been the case had one articulate professional user group been negotiating issues of structure and design.

Unified systems also offer the potential for increased centralization and control by library management. Again, I think libraries have been somewhat slow to exploit the ways in which computerized systems can contribute to managerial control; but I have no reason to doubt that that will happen. The economic pressures are too great not to do so. Greater managerial control is likely to result also from the deprofessionalization of the work force in which the professional management becomes increasingly distanced from nonprofessional catalogers and other workers.

**Distributed Systems**

Distributed catalog systems provide advantages to professional users similar to those provided to public users. Work can be carried out at different locations, not only in various library locations but also at home. Several individuals can work on the same database at the same time. Distributed systems provide important access to cataloging information to professionals in all areas of the library, and thus contribute to coordination and unification of delivery of service.

A major unintended consequence of the development of such systems is the separation of public and professional user that was briefly discussed earlier. Although much valuable information about public users can be gained from systematic analysis of online use, it is not the same kind of information about the public that is built up through face-to-face interaction. Without that interaction, it is easy to see users in terms of their familiarity with the online catalog (naive versus experienced users) or their approach (known-item versus subject searching). Information about the states of knowledge with which users approach the catalog or social or cultural differences that may affect the ways in which users approach the cataloged information is not readily obtainable through a secondary analysis of use of the online catalog.

**Conclusions**

To summarize briefly, users of online systems have the advantage of increased points of bibliographic and physical access to materials. Potentially, the online catalog can provide access to more recent materials more quickly; and, when it is used for interlibrary loan, the online catalog can more easily provide access to a wider range of materials. These gains are important in a society that tends to value speed, comprehension, and ease of access in its library services as well as in its department
stores, banks, and other organizations.

And the costs to libraries for such gains are similar to the costs to these other organizations. First, the work is less visible both to professionals and clients in the organization; it is less clear exactly how the numbers/words/commands get into and out of the machine. It, therefore, becomes more difficult to verify or check one's work or to feel in control. Second, there is increased isolation of users both from the organization when using online systems at remote locations and from one another. Third, substitution of computerized for manual systems provides increased opportunities for managerial control over the pace and nature of work.

I make this analogy between libraries, banks, supermarkets, and other organizations intentionally; for although the psychological and social effects of libraries' adopting online public access catalogs are potentially similar to the effects of adoption of computer technologies in other types of organizations, the overall effects on the organization are much greater in libraries. Information is not a commodity like Tide or Breyer's ice cream. Nor is a library's providing information to a user a transaction like cashing a check or opening a savings account. Access to information, I would argue, is in part a function of the social context in which it is given and received. Although those of us who talk about marketing library services refer to those services as "products," they are not quite the same.

One of the major focuses of our profession for many years has been training and research in information use. We have come to believe—and I don't see that belief as some self-serving ideology—that the extent to which information is appropriate to public users' needs is in part dependent on human professional skills. Collection development, "the reference interview," and various types of public services are geared toward bringing professional skills to bear on the central problem of how people obtain the information they want and need. Our struggles with subject access are just one indicator of the problem of trying to systematize user access. There are others.

Research by Robert Oddy, now at Syracuse's School of Information Studies, and others has demonstrated that successful retrieval of information is dependent on an understanding (or at least matching) of the user's mental representations of a query.

It should also be noted that the information access problems that are solved by online public access catalogs are problems for only a select body of users—principally academics and certain professionals. What about those users who want only one or two good books? What about those users for whom the value of time is different? I suspect that for many library users the value of information does not rest with the qualities provided by the installation of an online catalog.

This is not to argue against online catalogs. It is not a Luddite attack. I want to suggest several things by this analysis.

First, we should be clear about our reasons for moving toward implementation of online public access catalogs.

1. To what extent are economic pressures on libraries—particularly
increasing labor costs—encouraging technological—rather than labor-intensive solutions?

2. Are organization goals of efficiency, productivity, and economy that suggest technological solutions in conflict with public user goals for service?

3. Are the problems in access that are solved by the online catalog the major problems encountered by users in your library?

4. To what extent are market pressures—pressures from the vendors, bibliographic utilities, library systems—affecting the technological decisions of libraries?

5. Once a commitment has been made to the fixed costs of computer and catalog maintenance, what options do library organizations have when faced with financial problems?

Those individuals and groups responsible for the design and implementation of online catalogs have spent a great deal of effort studying users of both types. Engineering studies have been conducted to determine the best angle of the screen, how intense the display should be, how fast it should move, what graphics are useful, etc. Psychological studies have examined resistance to change, user anxiety, and fear of machines. The Council on Library Resources Online Public Access Catalog Study has asked questions about individual use of and satisfaction toward the online catalog. In all of this work, however, users continue to be treated as isolated individuals—not as social beings. I gather that the second round of user studies for the online catalog will try to fill that gap somewhat by examining more directly some socio-demographic aspects of users and nonusers. But there are broader questions still. For example:

What are the effects of online catalog use on the way in which people think about and use information?

What are the effects on the internal organization and the library profession at large of the changing structure of professional work in libraries?

How will the changing structure affect the development and delivery of library services?

We must ask these types of questions for technological decisions are choices not inevitabilities. The choices we make have significant and permanent consequences for public access to information and for the structure of the library profession.

REFERENCES


2. Phyllis Jayne, personal communication with the author. Dartmouth College Library is designing a link between their circulation system and a delivery system, according to Phyllis Jayne, director of user services, Dartmouth College Library. They expect to be operational within two years.

Use of Classification in Online Retrieval

Elaine Svenonius

Sometimes it is argued that with the automation of information retrieval it is possible to dispense with traditional methodologies for organizing information, in particular, classification. Perhaps the strongest counterargument to this is that classification underlies all thinking; thus it would be prima facie surprising if it found no place in online systems of the future. But what is this place? A difficulty in answering this question is that often our view of the future is obstructed by what is happening immediately around us. Nevertheless, it seems important to attempt to sketch some scenarios depicting how in the future classification might be used in online systems.

One way in which classification can be used in online systems is to increase the number of relevant documents retrieved by broadening a search. Suppose we are interested in protein requirements for women and, in keying in the search terms protein and women, find that we retrieve very few relevant documents. We could broaden our search by keying in terms that include our original search terms. In the BIOSIS database, a term immediately broader to protein is Nutrition—Protein, Peptides, Amino Acids. There are different mechanisms for broadening a search. Where a classification is expressive, i.e., where a classification notation reflects or expresses hierarchy, a search can be broadened by successively dropping the right-most digits of a classification number. Where a classification is not expressive, or not wholly expressive, broadening a search requires assigning to each document, manually or automatically, a chain of class numbers of different levels of generality. In effect this is what is done by the use of category codes in some of the online databases.

This is the method by which classification can be used online to improve recall, i.e., the retrieval of relevant documents. That classification can be used in this manner has been observed by many. Not so often observed, however, is that classification can also be used to improve precision, i.e., to screen out unwanted documents. Suppose we keyed in the term Nutrition—Protein, Peptides, Amino Acids (or the code corresponding to that term) and women, and found that most of the citations retrieved...
dealt with pathological cases, for instance, with a protein deficiency associated with some illness. This result presents a problem. To key in the additional search term *healthy* would probably not do much good, since it is unlikely that a document about healthy women would ever include mention of the word *healthy.* *Healthy* is a general concept and as such its mode of expression is various and unpredictable. But coming back to the problem of how to improve precision: a possible approach would be to enumerate, for exclusion, all pathological states. That is, we could say that we wanted everything on *protein* and *women* and NOT (*diabetes* and *pneumonia* and *tuberculosis* and etc., etc.). But we would hardly have time (or imagination) to key in every conceivable pathological state in the parentheses following the NOT operator. Our task would be much simpler if the database incorporated a certain amount of classificatory structure, so that every document assigned a term designating a specific illness were also tagged to indicate Pathology—General States. We could then search on *women* and *protein* and NOT (Pathology—General States). The point here is that the only efficient way to achieve desired precision in retrieval from a database may be to exclude from an initial retrieved set of citations a whole class of items. This approach presupposes that the database be classified.

Classification in online retrieval may be used to improve recall and to improve precision and—not to be overlooked—to save time at the terminal in keying in. But there is a caveat. Only in some areas can classification be exploited in such a manner. Earlier it was mentioned that classification was basic to thought; but it would seem there are some areas of knowledge where it is more basic to thought than others. Biology is one such area. Not only retrieval of information but also research in an area such as biology presupposes well-developed systematic taxonomies. Living things are divided into animals, plants, etc.; animals are divided into *Homo sapiens,* etc.; *Homo sapiens* into male, female, etc. Taxonomic classifications such as these are pure or logical in the sense that every member of a subclass is also a member of the class(es) including it. When it is said that a classified order is "logical," generally what is being referred to is the logic that governs the inclusion relation. A mathematical expression of this logic is the statement above about membership; every member of a subclass is also a member of the class(es) including it. Every member of the class of females is a member of the class of *Homo sapiens,* which in turn is a member of the class of animals. The picture is of a circle within a circle within a circle. But not all knowledge admits of neat taxonomies.

Library classifications are not always neat. In fact, in many the logic defining class inclusion is violated. To take Fairthorne's "excruciating" example, the class of *slide rules* in the Universal Decimal Classification is included in the class of *calculating and adding apparatus,* which in turn is included in the class of *apparatus with wheel mechanisms.* Thus, an attempt to improve recall by broadening a search on *slide rules* would retrieve all citations dealing with apparatuses incorporating wheel mechanisms. Another example, this one from the Dewey Decimal Classification (DDC), is the *abominable snow man,* where the inclusive chain of classes is...
first mysteries, then controversial knowledge, and then knowledge. It is hard to imagine a user who would want to broaden a search using a chain like this.

The questionable logic of library classifications is even more apparent when, in going beyond natural phenomena, they attempt to classify concepts or ideas. How are ideas such as love, truth, and beauty to be classified? This difficulty was treated as a serious problem by Plato who wanted to understand how such ideas “participated in each other,” i.e., how they were connected or related. In more recent times, the logician Jevons dismissed the problem declaring library classifications to be a logical absurdity. Somewhat later Kaiser simply refused to classify abstract ideas and in his Systematic Indexing recognized only concretes, processes, and geographical regions. The fact that some areas of knowledge are not so susceptible to hierarchical structuring as others is of particular relevance when considering the uses of classification in online retrieval.

The distinction between things that are relatively easy to classify, such as natural phenomena, and those which are not, is implied in a distinction that used to be made between natural and artificial classifications, that is, between classifications designed to order objects in the real world and those designed to order books on shelves. This distinction might be looked at in a way that better illustrates how classification may be used online, viz., as a distinction between taxonomic classifications and perspective classifications. Taxonomic classifications are those that are logically pure and, thus, can be used as described earlier to improve recall and precision. Perspective classifications, on the other hand, are not particularly suited to these uses, but they can be used in other ways. They can be used to focus a search (1) by contextualizing search terms and (2) by enabling browsing.

A perspective classification can be used to contextualize search terms. Consider a reference librarian who is trained to “negotiate” questions. Part of this negotiation is linguistic in nature; that is, it takes the form of contextualizing a patron’s search terms. Suppose, for instance, that a patron approaches the reference desk saying that he wants something on freedom. How would a reference librarian react to this request? Probably the first step in a negotiation would be an attempt to pinpoint meaning: freedom of what? —of thought, of will, of the press, of speech . . . ? In the designing of user-friendly online systems, the reference librarian’s activity of pinpointing meaning can be in part simulated by the use of traditional classifications such as DDC. For instance, a patron (in the online situation called a “user”) keys in the word freedom. The computer’s first step is to consult the DDC relative index, where it finds that freedom occurs in a number of phrases: freedom of choice, freedom of information, freedom of thought, freedom of the press, freedom of speech, etc. On being asked to choose from among them, the user may select freedom of choice. Again the computer peruses the DDC relative index and determines that freedom of choice is treated from two different points of view: from the point of view or perspective of Christianity and from that of natural religion. Again the user is asked to make a choice; he may choose Christianity. For the third time the computer consults the DDC relative index this time to
find where in the schedules freedom of choice in the context of Christianity is treated. It is found at the hierarchical address 233.7. By going to this address the computer can gather the information needed to display to the user the hierarchy of meaning embedded in 233.7:

200 Religion
230 Christian theology
233 Humankind
233.7 Freedom of choice between good and evil

Thus by contextualizing vague words, such as freedom, within perspective hierarchies, the computer might guide a user from an ineptly or imprecisely articulated search request to one that is quite specific.

Another way classification can aid in focusing a search is by enabling browsing. Browsing capabilities have already been implemented on some systems, for instance at Ohio State University. On that system the user can ask to browse through a number of items in classified sequence on either side of a given item. Scanning bibliographic citations arranged in a classified sequence, though not quite like browsing among books on library shelves, can be an effective way to focus a search. Indeed, the fact that many bibliographies (for instance, the British National Bibliography) are published using a classified order indicates that useful browsing need not be confined to library shelves. If online browsing is to be a feature of catalogs of the future—surely it must be—then some provision must be made for a classified ordering of citations. While an alphabetical ordering supplemented by a syndetic structure affects a certain degree of collocation even as Cutter observed, it can do so only limitedly. A question of some importance is what kind of classification is best suited for online browsing? Faceted classifications and fragmented polyhierarchical structures, such as those used in thesauri, are in vogue now; however, it may be that classical schemes of the enumerative type would better accommodate online browsing since browsing, whether done on- or offline, usually incorporates linear scanning. In any case, it would seem that effective online browsing presupposes a classified database. Online catalogs appear to have awakened a new interest in subject access. Perhaps we may see in this country an awakened interest in the classified catalog.

Several other uses of classification in future online systems might be mentioned briefly. One is the use of classification in the design of nonbibliographical databases. Travis points out that improving the present quality of statistical data, for purposes of reporting as well as retrieving, requires the rigorous definitional framework that only a classification scheme can provide. Another is the use of classification to provide collocation of a kind heretofore not possible in manual systems, viz., the collocation of documents that are like each other by virtue of sharing linguistic features in common such as similar index terms or similar citations or a similar natural language vocabulary.

A final and particularly important use of classification in online retrieval systems is as a switching language. Considerable interest today is being shown in the compatibility of retrieval languages. This interest is
perhaps greater in Europe than in this country since compatibility is of special concern where the databases to be searched are multilingual. A classification such as DDC that has already been translated into many languages comes ready-made as a switching mechanism. With the use of such a classification, a user’s search terms entered into the system in one language can be switched through DDC numbers to retrieve documents in several different languages. To develop a new switching mechanism, such as the Broad System of Ordering, rather than to use ones already existing, seems inappropriate; as inappropriate, say, as to develop a new language such as Esperanto. It can be an asset for a classificatory language to have the backing of tradition; it is certainly an asset for it to have financial backing.

In summary, eight uses of classification in online retrieval systems of the future may be envisioned. In areas of knowledge admitting of natural taxonomies, classification can be used to improve recall and precision and to save the time of the user in keying in search terms. In other areas of knowledge, perspective hierarchies can be used to contextualize the meaning of vague search terms, enabling the computer to simulate in part the negotiation of a search request carried on by reference librarians. An important use of traditional classification in online systems is to provide a structure for meaningful browsing. Classification can be used to provide a framework for the representation and retrieval of nonbibliographic information, e.g., statistical data. Automatic classification can be used to collocate citations in ways not possible in manual systems, e.g., by similarity of linguistic features. Finally, classification can be used to achieve compatibility of retrieval languages by serving as a mediating or switching language. In short, not only does classification have a place in online systems of the future; it likely has such an important place that we should prepare for a resurgence of interest in both its theory and its practice.

References

3. This example was shown to the author by Anthony C. Foskett.
The practical considerations of subject access addressed in this paper can be put as three questions: (1) If we have keyword searching, do we need subject cataloging also? (2) How much credence do we give to the users' views of the online catalogs? and (3) What is wrong with the concept of user-friendly systems?

To begin with the question of keyword searching versus subject cataloging—we’ve all experienced frustrations with the inadequacy of the subject catalog when looking for material in card catalogs. And we all know how expensive it is to do this inadequate subject cataloging. So when we were introduced to the concept of keyword searching—years ago, you’ll remember, computer systems people taught us about Boolean logic, KWIC and KWOC, full-text searching, and indexing—some librarians thought that the computer would solve the two problems of the inadequacy and expense of subject cataloging.

With the ability to search by keyword, when a user asks for works on a subject—for example, Chicanas (feminine version of Chicano) or birth defects, terms for which there are no LC subject headings yet—the online catalog displays citations with those keywords in the titles. Put that together with the new wisdom that 99 percent of all nonfiction titles use words that describe the contents of the book, and you might be convinced that there is no more need for subject cataloging. WRONG. Don’t for a minute believe it. While a keyword approach is useful and allows the user a flexibility in searching not possible in the card catalogs, assigned subject headings are as essential as ever.

There are severe limitations to keyword searching. Just as the authorized heading approach doesn’t pick up new terminology, the keyword approach doesn’t pull together titles that use obsolete terminology.

The intellectual organization of the catalog must still be maintained by human beings who assign subject headings that link and integrate ideas to the bibliographic record using a controlled vocabulary. If we
abandon this principle, the entire burden is on the users to pull materials on a subject together, to think of synonyms, including equivalents in foreign languages—and even then they'd miss all the books whose titles don't give a clue about the subject.

So, for example, in the University of California online catalog serving the nine-campus system, which is known as MELVYL and is still in experimental form, when the user searches by subject in the simple version of the catalog designed for new users, the system responds with citations that have the word or words anywhere in the subject heading or in the title. To the extent that this method produces irrelevant or too few citations, the user can then look at the subject headings assigned to those relevant records retrieved from title keywords, and search on those subject headings—e.g., a record with the word “Chicana” in the title might have a subject heading Mexican Americans, which the user could then use to search further.

And while we’re in a skeptical mood, we can also throw out all those articles that were written in the not-too-distant past reporting studies that proved conclusively that 90 percent of all library searches were for known items. These studies lent support to the idea that we could save a lot of money by stopping subject cataloging. You might conclude too from these findings that, in a pinch, even keyword access could go.

There is a lot wrong with that reasoning. First, subject access (by keyword or authorized subject heading) even in card catalogs is an important way to get at a known item when the patron is using an incorrect citation. After all, the patron who knows the book desired hasn’t necessarily seen the title page of that book. Chances are pretty good that the patron will be looking under a misspelling, a wrongly remembered title, or a correctly remembered title to which the library has provided access only by a symposium or a corporate author or some other unfriendly main entry.

Also, I suspect that many library patrons come to the card catalog with a known-item search because they think they have to. They don’t understand the idea of subject access except by going directly to the shelves.

The Council on Library Resources’ Online Public Access Catalog Project to solicit users’ evaluations by an extensive questionnaire—now in progress on a mammoth scale—is showing that subject access may be becoming more popular than previous studies indicated. This especially seems to be true in systems allowing access by keywords within subject headings or titles. Observers of people using such systems are finding that people looking for known items by keywords are happy to learn about other material on the same subject that is discovered in the process. We shouldn’t be surprised if future studies find that subject access is chosen more often than author/title.

On the second question, how much of what users tell us to believe, I am finding some interesting things. First, users love anything that is online—regardless of the system. The younger ones take to it instantly, the older ones are won over the moment they put their toe in the water. (The trick is to get them to put their toe in.)
But that enthusiasm lasts a relatively short time. The novelty wears thin, and as the user becomes seasoned, frustrations set in and they become impatient with the limitations of the catalog. To the extent the users feel it is their fault they are not finding things, they go away without asking for help, just as they do when they don't find things in the card catalog. If they are aware that the online catalog makes searching more difficult for them, they go away angry and impatient.

I don't mean to say that you will ever go back to cards once you are online: but so far, in evaluating today's online catalogs, it is easy to be fooled by patrons' enthusiasm into thinking (a) they are using the catalogs properly or (b) they are getting what they want.

It's easy to see why online is liked.

The online catalog holds the same magnetism, fun, and challenge as the thousands of games with which people five years old and up spend millions of hours. The principal reason it is liked is its interactive nature. The catalog doesn't just sit there, obliging the user to think of the next step; the online catalog talks back, makes suggestions. It gives the user a sense of control, a feeling of personal attention.

Also, there are many questions that can be answered by existing online catalogs that could not be answered by the card catalog.

At the Evanston Public Library in Illinois, an eighteen-branch library system, patrons in any one branch can, for the first time, have up-to-date information about the circulation of copies in any of the other branches, with the user-access mode of the CLSI circulation system on touch terminals. Theoretically, there's no limit to how far you can decentralize. On my campus, many departments use terminals they already own to dial our catalog.

At Lansing Community College in Michigan, which also has the CLSI catalog, using both touch and keyboard terminals, the user has access to on-order information for the first time. Librarians there especially like the benefit, too, of no filing backlogs. In some places, the absence of filing backlogs has created a new problem: an item shows up in the online catalog before the librarians have had time to shelve the book.

A librarian at Stephen Austin State University in Nacogdoches, Texas, which has a DataPhase public catalog, notes that searching for alternative subject headings is much easier in online catalogs than in card catalogs.

So far, these online public catalogs based on the circulation systems I've mentioned do not provide keyword access. You must enter the subject heading exactly as LC lists it. MELVYL, the catalog used by the nine-campus University of California system, provides quite sophisticated access: keyword in title or in subject heading; ability to limit a search by date, or campus, or group of campuses; two modes of use: one for new users, which guides them through the search step by step, asking them to make choices from lists and the other for experienced users, which is faster, more flexible, and capable of handling many more kinds and complexities of search problems.

These are all very good reasons to be enthusiastic about online catalogs. They do save time. They do allow questions to be answered that
wouldn’t have been possible with cards.

But don’t conclude much from the fact that the users say they like it, even love it. Don’t conclude that because the user feels friendly toward the online catalog, that the catalog feels friendly toward the user.

The online catalog can be as uncommunicative and as unhelpful as the card catalog ever was.

What are some of the things that upset users—and should upset us? All of the online catalogs with which I am acquainted tell the user who misspells, even slightly, or adds a wrong punctuation mark or adds an extra space, or takes one away—or what is so very common—enters the singular form of the word when the catalog has it only in its plural form—all systems in my experience tell the user who has committed one or more of these sins that the desired entry isn’t there. They are all unforgiving, to one degree or another, of user habits—habits that weren’t fatal in the use of the card catalog because the user could see an inclusive range not allowed in the computerized catalog.

That limitation is not inevitable. Systems can be designed that make allowances for the situations I’ve described. Taken together, perhaps all of the systems today make allowances for all of those situations—in different combinations—but no single one goes the whole way.

Also all of the catalogs with which I am acquainted use jargon: “Index,” “No match found.” The term call number is jargon. How else to explain why, in the Council of Library Resources user evaluation project, a number of people using my library catalog said “Easy” in answer to the question, “On a scale of 1 to 5, is it easy or difficult to search by call no.?” That’s especially peculiar because you can’t search by call number on my catalog. They all require the user to understand the difference between an author and a title and a subject. That’s knowledge that doesn’t come naturally, and something you didn’t need to worry about in card catalogs (at least dictionary card catalogs). Frankly, I don’t know how easy it will be to get users to understand the difference. Recently I was asked by an English professor at the University of California library why we had no books about the author Petr Kropotkin. The catalog said we had books only by Kropotkin. It turns out that this professor, who we must assume is a very smart man in all other respects, told the computer he wanted “author Kropotkin” when asked if he wanted author/title or subject. He didn’t understand that sometimes an author is an author, sometimes an author is a title, and sometimes an author is a subject, and he needs to know which is which when talking to the computer.

Another problem that all of us are aware of is the blind faith the user has that the computerized catalog is all-knowing, all-telling. Most of us do not have our old catalogs fully converted. At York University in Toronto, for example, where they use a public query subset of the Geac circulation system, not all of the material in the online catalog has subject headings; but that is not understood by the user, who thinks that the computer is telling the whole story when asked for material on a subject.

In using MELVYL, which at present has nearly eight hundred thousand records of holdings on all nine campuses, patrons do not under-
stand that when other of the campus libraries are listed as holding a title that interests them but Berkeley is not, they may not have the whole story. Even though the first screen of the online catalog warns “This is NOT a complete record of books in the UC libraries. Check your local catalog,” they come trotting to our interlibrary loan office without searching the title in the Berkeley catalog. Some of us have to come near to fistfights with our patrons, trying to convince them that the computer doesn’t know everything. As if the computer were an infallible mind, the patron insists, “But the computer told me . . .” or “I asked it for X and it told me thus and so.” With the card catalog, the verbs are quite different. You don’t ask the card catalog anything, you look things up in it. The card catalog doesn’t tell you anything; you read thus and so on the card.

After a while the users catch on to the fact that there are great limitations to the online catalog. They don’t know how to get out of loops they’re stuck in; they are sure a certain book is in the library, but the computer is telling them it isn’t. Rather than ask for help, they go back to their old friend the card catalog if it’s still around—which is no longer such a friend because usually we have stopped keeping it up to date.

We have a long way to go before we really understand the user-terminal-catalog interface. One thing is certain: if there is a way to make a mistake, to misinterpret instructions, there will be a large number of users who will find that way. Many wonderful stories are emerging about perfectly understandable misinterpretations by users: one catalog gives you a list of options to choose from, then says “type your choice”; so some users type “your choice.” I found a man the other day typing “return” because my catalog instructs you to press return after keying your side of the transaction. He was indeed pressing R-E-T-U-R-N; he didn’t realize there was a return key. To see a previous screen in MELVYL, there’s a line at the bottom that says TYPE PS TO VIEW THE PREVIOUS SCREEN. A great many people don’t read instructions, or else they don’t notice them under certain conditions, so we often get the question, “HOW DO I SEE THE PREVIOUS SCREEN?”

One great improvement over the card catalog in online terminals that use keyboards is that the user doesn’t need to know the order of the alphabet. I think we have not realized what a problem that can be until the touch terminals came along.* Several librarians have mentioned to me the user frustrations over not being sure whether “P” comes before or after “R”; “L” before or after “J”—knowledge you must have in using touch terminals to answer such a question as “Does your word fall between brain and branch? Brand and brazen? Brazil and breakfast?” Some librarians have pasted the alphabet onto the terminals to help solve that problem.

To summarize this point, don’t be disappointed if user enthusiasm turns sour; there are good reasons why that could happen.

Turning to my third point: the question of user-friendly versus user-hostile systems. We’ve put a great deal of faith in the interactive ability

*Editor’s note: For more details, see p.43–47.
of the computer—that is, its ability to respond to the particular thoughts of the user as keyed on a terminal and to conduct a kind of dialogue; we’ve put a lot of faith in that power to solve what we saw was a great problem with users’ lack of understanding of our card catalog. With cards, we can’t be there whenever the user misuses the catalogs, makes wrong assumptions, doesn’t know alternatives. We’ve been hoping the computer could substitute for us, that there would be instant instruction, instant correction.

I’ve come to think that is not possible and even it were, not desirable. I’ve come to the conclusion that it is quite possible to design a friendly and easy-to-use system for very simple functions: a simple author or title or subject search. But I think it would be wrong to expect users to learn on their own how to perform the complex searches that can be performed infinitely better in a computer catalog than in a card catalog.

Since making the system user-friendly means that the patron is always presented with a manageable set of choices, and since you cannot write a set of choices for every conceivable situation, “user-friendly” to me is synonymous with limited service.

I don’t want to imply that every system that provides three choices is always user-friendly. But it is certain that every system that provides one hundred choices at one time is definitely not friendly.

Every librarian with online catalogs to whom I’ve spoken agrees that “library instruction” has taken on a new meaning. Users are more receptive to being taught about the catalogs than they ever were before; we are more aggressive than ever before in devising clever ways of reaching users.

I think as online catalogs become more sophisticated, we’ll see that in-person instruction is more necessary than ever, not less. We won’t be able to rely on online instruction, though it should be there. We won’t be able to rely on signs and brochures, though they should be there—too many people don’t read them.

We should insist that every online catalog provide easy access for simple searches by first-time users or those who want to perform simple searches. Such a catalog version should be self-explanatory, it should be quick, it should inform the user when he is making a mistake and tell him how to correct it; it should take no learning time. It should be VERY user-friendly.

This access for simple users should not force the user to do things that don’t fall within common sense or common practice. At present, our so-called simple online catalogs are far too rigid. To use the DataPhase catalog successfully, for example, you are not allowed to type periods after personal-name initials. You are not allowed to type in punctuation when searching a corporate name. You’ll get a response of “no match found.” With the CLSI catalog, if you don’t use the truncation symbol, you’ll miss variant versions of an author’s name, including those where the only difference is dates or other designation after the name. Or if there’s nothing in the system matching just what you’ve keyed, the system reports “not on file” when in fact there may well be lots on file. With the MELVYL system, when you try to narrow a search, sometimes it lets
you, other times it doesn’t—but common sense wouldn’t tell you why. To the user, there is no difference in the kinds of operations he or she is asking the catalog to perform; to the catalog there are big differences.

So for simple searches, the catalog system should be designed for the worst case user.

But when it comes to more sophisticated capabilities, we should not look for simplicity, or cordiality. Of course, whatever language is used to instruct or direct should be clear and free of jargon; the language itself should not be hostile (one system says “illegal command”), but it need not be simple. The catalog must, however, make it perfectly clear to the user that if he or she is not inclined to take the time to learn, the “advanced mode” cannot be used correctly, though there is always the option of using the “simple mode.” Oddly enough, it may be easier to get that point across by creating a larger gap in the ease of use between the “simple” and “advanced” versions of the catalog—though I would not want to press that point too far.

To close, I want to leave you with three basic messages: keep both keyword and subject-heading access in online catalogs; don’t put too much stock in the user’s uncritical attitude; there are many reasons to be critical of online catalogs; and finally, think about whether your “user-friendly” systems are falsely lulling users into believing they understand the catalog when, in fact, they don’t.

All of the online catalogs I’ve mentioned are in the midst of change and development. We have an opportunity to shape them. That’s a very exciting opportunity, and one that may not last forever. We shouldn’t let it slip by.

**Reference**

Futuristic Aspects of Subject Access

Phyllis A. Richmond

The future, as the cliché goes, lies ahead. It almost certainly contains a computer in it for each of you. All kinds of things have been forecast for what this ubiquitous machine will do to and for you. Some of it sounds very interesting indeed. This future will be discussed in this paper first in terms of the kinds of hardware that are coming along, then in terms of the software, next in terms of "friendliness," and finally with regard to what can be done with it for subject access.

The computers themselves will continue to be very small in size, with the density of circuits per quarter-inch-square increasing steadily to some ultimate end in smallness, not yet reached. At the same time it will become possible to do more and more with this very tiny integrated circuitry. How much more it will be able to do than is being done now remains to be seen. Where it may now take several discrete chips to form a computer, in the future a single one may do everything. At the same time the prices constantly decrease.

The amount of memory "on board"—i.e., immediately available for use—has increased steadily, a development that means you can do more than ever quickly. Twenty years ago, strong men turned blue when one wanted to work with words—our kind of words—because there was not enough space for the computer’s version of our kind of words. Since then, the size of the chunk that the computer can handle at one fell swoop has increased in geometric progression and, therefore, the computer can work on the data much faster. The combination of more and faster is just what we wish to see in online searching in various combinations of indexing and classification.

Some other things also are taking place. A cheap computer is available now for $99.95. It can do a remarkable number of things, considering its size, and it is designed to be reasonably "friendly" to the user. For one thing, it will not let you write a line of a program unless it is done correctly. Because of this feature, if your reasoning has been correct, the completed program should do what you wish without giving you a lot of

---

Paper presented at the RTSD/CCS Subject Analysis Committee program on "Subject Analysis in the Online Environment" on July 13, 1982, by Phyllis A. Richmond, professor and acting dean, Baxter School of Information and Library Science, Case Western Reserve University.
nasty error statements that require corrections.

The big computer companies, IBM and DEC, have joined the group and come forth with personal computers that are one jump ahead of the kind like Apple that began in someone's garage. In Britain, the BBC has chosen a microcomputer for mass production, to be attached to the home television set. Various kinds of specialized computers are available for business, for scientific work, for graphics in color, and so on. Operating systems are improving. It is not unlikely that the QWERTY keyboard will be replaced with something easier for the nontypist to handle. Now available are computers that can hear and answer or ones that enable you to communicate by just touching the face of the cathode-ray-tube monitor screen.

If people haven't rushed out and bought personal computers by the millions, it may well be because there is always an improved one that will do more waiting in the wings. Presumably there is an end to the constant parade of new microcomputers, but it is nowhere in sight yet. All these advances mean that today you can get a tiny computer to use at home that will do more than computers in common use could do in the 1960s when library automation began. It is becoming possible for small libraries, for library schools, and for librarians to acquire their own machines. Since these computers can be adapted to library needs, we no longer need be at the mercy of programmers who pale at the merest thought of working with words instead of numbers, equations, and formulas.

Along with the computer itself, peripherals are becoming more and more intelligent. The cathode-ray-tube monitor that hangs on the wall (instead of requiring the space of the standard TV) is within sight. It has come to the point where an artist can paint with light on the face of the TV tube, as was demonstrated at the annual meeting of the American Society for Information Science in Washington in October 1981. Computer music is now so commonplace that we hear it every day, probably without recognizing its source. It should not be too surprising someday to encounter one of those marvelous Library of Congress subject headings that will read “Suites for 2 violins, 1 oboe, 1 cello, 1 bass and 1 computer.”

On the software side, things are just as exciting as they are with the hardware, albeit taking place more slowly. For one thing, an ANSI standard for BASIC is in the works. If all goes well, it should be available by 1984. If the standard is followed, one will not have a different version of this computer language for every machine, as is the case now. When that happens, David Lien can finally stop updating his nice little handbook describing the variant BASICS.

A great deal has been written about “user friendliness.” A user-friendly computer is one that doesn't require one to have a Ph.D. in computer science to make it do what you want. The friendly computer helps you get what you are after without forcing you to understand the digital electronics upon which it is based. Some very fancy languages are appearing, less forbidding to common mortals. One is called LOGO, which is used to teach children mathematics without pain, something all
of us could use. Another, SMALLTALK, was developed at the Xerox Palo Alto Research Center. This one also was used for children in its earlier stages. It is of interest for subject analysis because it is based on a system of classification. Another version is called DYNABOOK, the name chosen in part because it was designed ultimately to be used with a computer the size of an ordinary book. Both SMALLTALK and DYNABOOK make use of the way the human mind recognizes patterns. It is thus related to research in cognitive psychology. In SMALLTALK there is a three-dimensional layering of patterns. From the surface the screen looks like pages, one behind another, which are connected in a classified system and may be called and recalled, for storing and accessing information in logically organized formats. This method may eventually replace the Boolean or other indexing strategies for online searching used at present. It is more like a thesaurus with logical connections between terms displayed physically. One would expect fewer or better defined connections than those based mainly on associative relationships. For this type, a study of the kind of classification found in the cross-reference structure is particularly apropos. The “propositional network” used by some cognitive psychologists to represent how the human memory functions is basically a classification system.

There are various kinds and levels of “user friendliness.” Nelson, for at least ten years, has been talking about a system called XANADU. Roughly, it is a system whereby one person thinks of something, communicates it via computer, and others add glosses to it, while still others add glosses to the glosses, and so on. Some groups are using this type of communication now, usually a group with limited membership, interested in some particular topic.

Recently, Pauline Cochrane, in *American Libraries*, described a system called PaperChase, “a computer program permitting users to search medical literature online without special training.” The system permits the user to form queries with an approximation of the term(s) desired. The computer has been programmed to accept a number of variants of an individual term. The user in need of instructions types a question mark and the computer responds with assistance, perhaps telling the user to try the root of the term for the disease or process of interest. When the user does that, the computer indicates how many references will be recovered and then displays some similar terms from a formal listing such as MeSH. Then it offers some options in menu form, and continues to aid the user until satisfaction results. The system is straightforward and economical. Finally the users are invited to add comments, and some of the improvements to PaperChase have resulted from this kind of feedback.

Another term one hears besides “user friendliness” is the word transparency. This term refers to a situation in which a query or an input of some kind results in a whole series of steps to satisfy the user’s objective, but none of the steps is apparent to the user—they all happen behind the scenes—and the result comes apparently effortlessly, just as if the computer had been sitting there waiting for someone to ask that specific
question or perform that exact action. Pournelle, for example, calls a
text editor named WRITE “truly transparent,” because one is scarcely
aware that it is operating—a helpful feature for the creative writer.12

One of the more interesting kinds of “friendliness” is a system called
THE LAST ONE, devised by two Englishmen.13 This is an automatic
program-code writer for the hard-pressed programmer. The program
writer does not have to learn a special language, but instead has to de-
velop a logical design for whatever he or she wants the computer to do to
some data. It is necessary for the user of THE LAST ONE to create a
logical design, similar to the Warnier-Orr diagrams in which one makes
a series of logical statements that are a progression of desired outcomes
from given inputs. THE LAST ONE is a tool, “a programmer’s design
aid,” which generates the actual code needed for the program.14 The hu-
man user has to define just what is needed in the proposed program, be-
ginning with the content of “files with their names and details of con-
tents.”15 Then the user proceeds to a description of what is to be done to
the files—the program proper. Menu-type prompts are fed to the user.
“As the user responds to requests for ever-finer detail, THE LAST
ONE builds a correspondingly more detailed model, until, at the end, a
complete description of the finished program exists in the form of a pre-
cise and complete ‘flowchart.’ ”16 Then THE LAST ONE structures
and codes this flowchart until, within a few minutes, it presents a fin-
ished program. If the user’s logic has been correct, the program runs the
first time. If it does not, like any other program, one looks for errors in
logic. The program can be corrected or augmented as needed; THE
LAST ONE fixes it up again. With such a helper, the busy librarian
might not even feel guilty about playing a few little games like Pac-Man.

Now what about our old bugbear: the problem of dealing with masses
of data. This is the stock in trade of anyone dealing with information.
There eventually will be lump or bulk transmission of data that have
been input and massaged into “ultracompactness.” You will be able to
carry around data on a small interactive videodisc, the content of which,
read by a handy computer, will do one of several things:
• read it to you
• print it for you
• use it for further operations you wish done

With recognition of your voice pattern as both a key and a lock—a key
to let you in and a lock to keep others out—you should be able to keep
your private affairs private simply by saying to your friendly computer,
“Open, O Sesame” or “Close, O Sesame.” It may even be possible to
protect programs with an instruction that says, “At this point stop and
accept only a voice command to proceed,” with that voice being yours.

The future looks much more inviting to the ordinary person who
wants to use a computer but does not want to spend a decade learning
how to do so. After all, some of the greatest inventions of the past have
been designed by and for lazy people. So all we need now is a mind-
reading machine! It would be helpful if the artificial-intelligence people,
instead of trying to create an intelligent machine, concentrated on how
to hook up a machine to our little gray cells, which we now know stay
with us much later in life than we ever suspected. This would enable us to capture some of those absolutely brilliant ideas we get in the middle of the night that somehow always seem to vanish in the cold light of dawn.

Finally, for subject analysis, the use of various new media promises an escape from problems caused by complete dependence on paper and film as our primary storage media. Like the PaperChase program, we can allow access via a wider range of entries—simply by letting the majority of users be successful with whatever indexing terms seem reasonable to them. Thus, every index term could be as global as classification and cross-references can make it. At the same time, with help from the cognitive psychologists and the artificial-intelligence researchers, we may be able to use our own heads to better effect in preparing our access queries.

It is my considered opinion, especially after participating in the Fourth International Study Conference on Classification Research in Augsburg recently, that there is not a very great future in the use of keywords, Boolean algebra, or LC subject headings in online searching. It is more likely that classification will be mandatory, perhaps of the type already begun by SMALLTALK. A news item in a recent issue of BYTE quotes the director of personal computer research at a marketing research firm called Strategic, Inc.: ""SMALLTALK will become the dominant operating system and programming language for 16-bit personal computer systems..." and "numerous leading-edge personal computer and office automation manufacturers will be flooding the market with SMALLTALK systems."" He anticipates that by 1983 Apple, DEC, Hewlett-Packard, Tektronix, and several Japanese companies will offer SMALLTALK versions, and IBM is probably developing a similar product. This sort of thing should greatly change methods of subject access and subject searching.

Some of the conferees at Augsburg reported that they use PRECIS because it supplies more possible searching points than are provided by keywords and Boolean combinations. PRECIS supplies strings of terms, with relationships indicated, both among the operators identifying single or multi-word terms themselves, and by the sequence (i.e., the string sequence) of these identified terms. PRECIS allows more places to look because the string strategy and the term strategy are built in as a part of its context-dependency aspect.

Meanwhile, back at the laboratory, the engineers and scientists are making friendlier computers. At the point where true friendship occurs, that is, when the friendly computer’s capabilities match the enhanced intelligence of the searcher, we may expect major advances in subject analysis. All in all, we should have a very exciting future.

REFERENCES

2. For information of this type, the column "Bytelines" in BYTE: The Small Systems Journal is informative. A similar column in DATAMATION is called "Look Ahead." Other journals have this type of column as well.
12. Jerry Pournelle, "'Terminal Madness, the Word, Grammatik, and Then Some,'" *BYTE*, 7:286-300 (June 1982).
13. "'The Last One'" (advertising brochure from DJ 'AI' Systems, Ltd. (Two Century Plaza, Suite 400, 2049 Century Park East, Los Angeles, CA 90067), p.[1-2].
15. Ibid., p. [2].
16. Ibid.
A New Information Access Tool for Children's Media

Theodore C. Hines and Lois Winkel

As a result of our work with the Children's Media Data Bank, we have accumulated considerable evidence of the practicability, feasibility, and need for a new type of subject-heading list or thesaurus to serve both as a searching tool and for vocabulary control in providing information access to children's materials. This paper presents some of the rationale for such a tool, together with detailed suggestions for its design, production, and use.

BACKGROUND

Information access to children's media is usually through the local catalogs of school library media centers and public libraries. These library catalogs are supplemented by a wide range of bibliographies and indexes, primarily analytical in nature (that is, they index bibliographic units contained within other bibliographic units: articles in journals, plays and poetry in collections, etc.). Many of these other tools use or adapt the techniques used for providing access through library catalogs.

Good library catalogs generally use the cataloging provided by the Library of Congress (including, of course, Cataloging in Publication [CIP]) or by a nonprofit or commercial cataloging agency. Despite the fact that the American Library Association has adopted Library of Congress annotated-card cataloging as the standard for children's media, most cataloging agencies (and most libraries) still use the Sears List of Subject Headings as their authority in providing subject access.

The number of cataloging access points provided for children's media
is comparatively restricted. The Library of Congress assigns an average of about 1.75 subject headings to each adult title and about three subject headings for annotated-card headings for each juvenile title. This average probably compares quite favorably with cataloging provided by other services or done locally by libraries. It may be also be compared, however, to the average of eleven descriptors for each document included in the ERIC (Educational Resources Information Center) database.

The relatively limited number of subject access points probably reflects the size limitations of the three-by-five catalog card and the catalog maintenance costs that would be involved if many more headings were provided. ERIC uses only selected headings in its printed indexes and retains the others in its machine-readable file for computer searching only. It would certainly be possible to follow such a policy with children's media at present, looking forward to the time when online catalogs and computer searching for children's materials would become common.

Searching for children's materials is now almost completely manual—so far as we can determine, only the Children's Media Data Bank has conducted regular computer searches for users—despite the fact that much of the required cataloging data are available in machine-readable form, either in Library of Congress MARC records or in the cataloging databases of various libraries and commercial firms.

While the information access provided for children's media is comparatively restricted, it is worth noting that the actual and potential audience for better access to such materials—children, parents, educators, and others who work with children—is very large in comparison to the audiences of many other databases. At the elementary and combined school level alone, there are more than 50,000 school library media centers in the United States, plus more than 14,000 public libraries including branches.

The social importance of providing appropriate access to materials to children during their formative years is evident. And, in view of recent discussions of networking, it should be noted that increased intellectual access to materials is at least as important as increased physical access. In fact, neither can function truly successfully without the other.

**THESAURI AND SUBJECT HEADING LISTS**

At this point it may be of value to introduce a brief review and comparison of information science thesauri and library subject-heading lists. One basic point has already been mentioned by inference. Thesauri are generally regarded not only as aids to indexing but also as basic tools for searching. Although knowledgeable librarians occasionally use heading lists as aids in searching, the lists are not basically designed for this purpose and their use in searching is neither usual nor usually known to library users.

Subject-heading lists generally consist of a single basic alphabet of headings, scope notes, subdivisions, cross-references, and reverse cross-references. They sometimes add also a brief separate listing of most generally applicable subdivisions.
Subdivisions as such are not usually a part of most thesauri, although they appear to a limited degree in some, *MeSH (Medical Subject Headings)*, for example. Thesauri, on the other hand, provide a multiplicity of listings. These may include a list of used terms, a major listing of all terms or descriptors and cross-references, a classified listing of used terms, a list of the hierarchies implied by the cross-referencing structure, and a permuted listing, generally providing access only to words in used terms.

Subject-heading lists may include classification for headings (as does the Sears list), but they do not provide access by the assigned class numbers. On the other hand, again with such notable exceptions as *MeSH*, the classification used with subject-heading lists is hierarchical and quite extensive as compared with the relatively brief and simple classification schemes used in most thesauri.

The cross-referencing structures used by thesauri and subject-heading lists are much more similar than their appearance suggests. *See* and *X* references in subject-heading lists correspond exactly to *USE* and *USED FOR* references in thesauri. *See* is perhaps clearer than *USE* and *USED FOR* than *X*. *See also* and *XX* correspond to *BT* (*BROADER TERM*) and *NT* (*NARROWER TERM*), respectively, except where they are reciprocal. In this latter case, they correspond to the *RT* (*RELATED TERM*) references in thesauri.

*BT*, *NT*, and *RT* appear much clearer than their equivalents in subject-heading lists, although the subject-heading conventions make better provision for what should appear in catalogs as cross-references for the user than the thesaurus conventions do for the printed indexes using them as authorities.

Hierarchical relationships seem much more rigorously indicated in thesauri than in subject-heading lists. In subject-heading lists they tend to be included only where they would be of value to the user rather than in every case where they are logically implied. The subject-heading list system may be the more valuable, as it permits more directly useful references to be included in a listing of a given size.

Also, the library subject-heading list cross-reference structure recognizes implicitly that a given term or heading may appropriately appear in multiple hierarchies at the same or different levels. This is a valuable insight, and is not—or at least not generally—true of thesauri, which appear to restrict term references to one level in one logical hierarchy of terms. Both systems may, of course, have any number of levels implicit in the syntactic structure. It is worth noting here that Charles Ammi Cutter, before the turn of the century, commented on the desirability of a listing of the hierarchies implicit in the cross-referencing structure of library catalogs, but felt that the task of creating such a list would be too difficult and expensive. Cutter might well have been pleased by the fact that a hierarchical list of this kind may now easily be computer-generated from the basic subject-heading list.

Library subject-heading lists, unlike many thesauri, do not place any restriction on the length of allowable headings, nor on complexity of form. Many thesauri do not allow inversions, for example, and some
prohibit punctuation altogether except for parentheses to indicate separate meanings of terms (an interesting exception).

An important comparative strength of subject-heading lists is their recognition of many types of headings as being implicitly included—names of species of birds, flowers, animals, etc., and place-names, for example—and that they may refer to external rules or authorities for the form of such terms: to cataloging rules for personal names, for instance, or to Chemical Abstracts for the forms of names of substances.

This procedure seems much preferable to some thesaurus conventions, where the term “identifiers” is used to comprehend both terms not yet admitted to the thesaurus and the type of entries referred to in library nomenclature as other added entries (for secondary authors, illustrators, or, in general, terms for things associated with but not the subject of the document in question).

Unfortunately, most subject-heading lists (and very many working catalogs) do not include cross-references to many of the most specific headings, using general cross-references (BIRDS See also . . . names of birds, e.g. Canaries) instead.

The limitations of library subject-heading lists may well be associated with the fact that they were developed in a precomputer era. Similarly, it might be said that most thesaurus development took place in a second-generation computer era, and that such thesaurus aspects as limitations on length of descriptors, lack of subdivision provision, and restrictions on the use of punctuation reflect the systems knowledge of the time and the capabilities of then-available equipment.

It is obviously possible to combine the strengths of both of these types of tools in the light of current knowledge and technology. This approach is the basic proposal of this paper.

**The Children's Media Data Bank**

The Children's Media Data Bank was the outgrowth of an earlier (1976-77) project at the University of North Carolina at Greensboro to investigate the applicability of computer-based information science techniques to children's materials. The project was funded by the U.S. Office of Education under a Library Research and Development grant. The first year's work suggested the desirability of starting and testing a prototype Children's Media Data Bank. Further funding was requested and received for this purpose and, still later, to transfer prototype data to microcomputers.

Both in order to keep costs down and because secondary-school children should be using a broad range of adult as well as juvenile materials, the data bank restricted itself to children's book trade (i.e., nontext) materials from the preschool through the sixth-grade levels. It has further limited the initial input to in-print titles in all media (except 16mm film) recommended by such basic sources as Booklist, the Children's Catalog, The Elementary School Library Collection, and School Library Journal. We estimate the total number of these recommended in-print titles at about 15,000.

At present the data bank covers approximately 10,000 titles. For
about six months we did intensive online and batch computer searches for users of children’s materials, working with a number of school districts, public libraries, and individuals. More than 500 searches have been carried out, many quite extensive. The programs have now been rewritten for microcomputer use, and a sample of the data moved to microcomputer disks for experimental purposes.

**RELEVANT PRODUCTS**

Products of this work relevant to this discussion have included a complete listing of the subject headings and subject headings with subdivisions assigned to the titles in the data bank, as well as frequency counts of the words in subject headings and the words included in annotations for all titles. Also of very great importance have been the insights arising from analysis of records of search requests and the resulting production of lists of terms used in searches and special classed lists of subject headings as searching aids.

**NUMBER OF SUBJECT HEADINGS**

As noted above, the data bank covers some 10,000 titles. The total number of different subject headings or subject headings with subdivisions assigned to the titles in the data bank is about 12,5000. This figure may be compared to the figure of about 15,000 subject headings or subject headings with subdivisions that were assigned to juvenile titles include in MARC through 1981.

It should be emphasized that these figures include all assigned subject headings or subject headings with subdivisions. Consequently, the number of subject headings proper is very much smaller than the totals given above.

**SIZE OF THE UNIVERSE**

The importance of these data is that they show that the size of the universe of subject headings and subject headings and subdivisions actually assigned to juvenile materials over, say, a ten-year period, is not only finite but relatively small. It should be noted that a very large number of headings included only implicitly in existing subject-heading lists, such as the names of individual person, places, and species of animals, appear explicitly in these listings. Included, for example, are such headings as the specific names of native American Indian tribes or groups, which are downward references from the heading INDIANS OF NORTH AMERICA.

The Library of Congress has not published a subject-heading list for children’s materials, although Lois Rose and Jane Marton, former and current heads of the Subject Cataloging Division, Children’s Literature Section, respectively, very kindly provided the information about the headings assigned to juvenile materials mentioned above. Instead, the Library of Congress has issued a listing indicating those headings used only with juvenile materials or used with special meanings when assigned to juvenile materials, together with an explanation of the policies used in applying headings to annotated-card titles.

It is not known how the Sears List of Subject Headings is compiled, or if
the compilation is based on the headings actually used in cataloging a collection or collections of materials, as should, of course, be the procedure followed.

What is indicated by our data and those provided by Lois Rose and Jane Marton is, however, that it would be quite feasible to base a new heading list on those headings actually assigned to materials, preferably on annotated cards by the Library of Congress as providing the national standard. If this were done on a rolling basis—that is, for example, including in each listing only those headings used during the last decade—it would be economically possible to include all of the headings actually used, down to the specific names of persons, species, substances, corporate bodies, and places. This procedure would not, of course, prevent the addition of new specific headings by catalogers as required. It would mean, however, that most of the specific headings needed as references would actually appear in the list, a very useful service to the cataloger but—and this is a critical point—a service to searchers not previously available.

From the point of view of economy of production of a new heading list or thesaurus, it should be noted that the headings and headings with subdivisions actually assigned might be derived from the MARC record without cost for keyboarding. Once an appropriate system had been established, only the cross-referencing and classification for new headings need be provided manually, and the new headings could be brought out automatically for the consideration of the editor.

**Types of Access to Children’s Media**

Conventional library cataloging practice provides various types of access to juvenile materials. In addition to main entry and series entry, the access points are generally divided into two groups: subject entries and other added entries. Other added entries include title, variant title(s), illustrators, editors, etc. The rules for added entries seem to permit the use of, broadly speaking, any term not a subject heading which might assist the user in locating the item; but this provision in the rules does not seem to have been as extensively taken advantage of as might have been useful.

Subject entries basically are intended to indicate the specific subject of the work taken as a whole. They may be subdivided in various ways, which will be discussed below. In cataloging generally and in the cataloging of juvenile materials in particular, “subject” entries are not actually limited to the subjects of books, but may on occasion include forms of materials (picture books, collections of essays or short stories) and genres of materials (mystery and detective stories, science fiction). They may also be used to indicate the geographic or historical setting of a work of fiction, and, in some instances, curricular relationships or possible uses of the material.

With certain exceptions in the case of the application of some types of what might be called non-subject subject headings to adult materials (collections of essays and short stories, for example), their use with juvenile materials seems to have developed from a sense that they were
needed in individual cases rather than from any theoretical basis. In some cases, notably in the instance of picture books and the various genre headings, they may have been adopted because the Dewey classification does not provide for the fact that many libraries arrange each of these groups separately on the shelves and need some cataloging indication of them. In practice this does not work out too well if such entries are actually included in the catalog, where they replicate an access order provided by the shelflist and are not really useful as primary access points.

It should be pointed out that the discussion above applies only to subject headings proper, not to subject subdivisions. Subject headings proper are generally intended to be specific; subdivisions, on the other hand, are designed generally to provide a useful classed breakdown of the titles assigned a given subject heading.

Based upon both recent and previous experience, we believe there is a need for access to subject subdivisions directly. This service cannot readily be provided through the catalog. It can, however, be provided to a significant degree through the type of thesaurus we propose. We feel also that there is a need to provide, more consciously, headings indicating forms, curricular relationships, and uses of materials; and that such headings should be distinguished from subject headings proper, perhaps by the use of upper- and lowercase letters.

**Basic Design of a Children's Media Thesaurus**

From the librarian's viewpoint, it is evident that many thesaurus techniques could be profitably applied to subject-heading lists, enhancing their value to catalogers and providing a new and valuable searching tool for librarians and library users.

We would like at this point to present an outline of the basic design of a children's media thesaurus intended as a searching aid as well as a cataloging tool. The data presented so far indicate that a list compiled on the basis of actually assigned headings over such a period as ten years would not be so large as to be impractical to compile and to publish. Our knowledge of Bradford and Zipf indicate not only that a list compiled on this basis would be likely to meet most of the needs of libraries, but that its size could be restricted, if necessary, by the elimination of low-frequency headings, particularly nonce terms not used in the past few years. These terms could, of course, be used and added to the list by the user quite easily if they should recur. They would also be retained centrally in a machine-readable database. It would seem highly likely that such a listing, despite being compiled for juvenile titles, would—like the Sears list—be useful for popular libraries in general, thus providing an even broader market.

Given the long period during which information science thesauri have been computer-produced, it is difficult to understand why the Sears list has not been automated. Based on our experience in compiling a number of computer-based thesauri, we can discuss the steps required, first to produce a kind of equivalent to the Sears list meeting current stan-
dards, and then to expand it into a true children’s media thesaurus.

We would assume that, in accordance with the American Library Association recommendation that Library of Congress annotated-card cataloging is the appropriate standard for school media centers and public library children’s services, the basic source of terminology and usage would be those headings assigned by the Library of Congress in that program. These headings, and their frequency of application, are available or can be made available from the MARC database already in machine-readable form.

The additional information required for a conventional type of heading list consists of the appropriate cross-references, scope notes, and Dewey numbers. The cross-references and scope notes can be derived directly from the Library of Congress subject headings and the Working List of Subject Headings for Children’s Literature. Editing would obviously be required.

Since there is such a close correspondence between Library of Congress headings and Sears headings as they are actually applied to books, it should be possible to derive a significant proportion of the Dewey numbers from the Sears list. If the Sears list is not current with Dewey at the time of compilation, then it may be necessary to classify the headings.

The labor required for the tasks indicated above is certainly quite significant, but the availability of the basic terms used in machine-readable form may make it much less than that which would otherwise be required. The use of thesaurus cross-referencing conventions in input would make it necessary to keyboard only one side of each reciprocal relationship—that is, cut the keyboarding of cross-references by 50 percent over conventional techniques. While we believe that thesaurus conventions (except for see references) are preferable (and would be preferred by library users), librarians we have consulted are very strongly in favor of traditional subject-heading references. Therefore, they should perhaps be retained except for Related Term (RT), since it would afford so significant a space saving in the thesaurus.

Further substantial savings in input keyboarding can be achieved by adopting the methods designed by Dr. James H. May using a simple abbreviation technique for high-frequency words or phrases, which are subsequently expanded automatically to their full forms. Also, obviously, there need be no re-keying of the data for printing. Indeed, given the nature of a subject-heading list, even greater economies might be achieved by using offset from direct computer output at a reasonable reduction ratio, since information science thesauri demonstrate that many of the typographic distinctions in the Library of Congress subject headings are not really required for efficient use of the list. Consideration might also be given, especially if the lists are expanded as indicated below, to producing the list or appropriate parts of it by COM, that is, directly by computer to microform. Between editions, listings of new terms might be published in serial form.

Obviously, the production of the second and all succeeding editions will be far cheaper than by present means.
Thus far, we have considered only the production of a quite conventional subject-heading list, with a main listing and a supplemental listing of most-used subdivisions. This automation of the list alone would be well worthwhile. We would suggest, however, including downward cross-references to all of the highly specific headings (names of species, etc.) actually used. We would also suggest that, whether they are printed out in the published list or not, the actual subdivisions applied to headings in the chosen section of the database be retained in the basic machine-readable source for the list.

A further suggestion is that in recording the data, personal names be tagged internally in some way (in our experimental work we simply used an asterisk), thus allowing for production of a listing of biographers. For both geographic headings and geographic subdivisions the indirect forms should be included and appropriately tagged where direct subdivision is used, e.g., <FRANCE>—AVIGNON and CLINICS—<MASSACHUSETTS>—BOSTON. Uses for this tagging are discussed below.

Once what might be described as the basic subject-heading list has been produced in machine-readable form, it is evident that a great deal of additional searching access and cataloging assistance can be provided simply by computer manipulation of the basic data.

Before proceeding to discuss additional listings, however, it might be useful to point out another way in which it would be possible both to reduce input keyboarding and at the same time provide more access than is usual in subject-heading lists. It is not necessary to keyboard any see references to inversions (or, to coin a term, reversions) of multiple-word headings. These can be produced automatically, using a stoplist or stoplists like those used in title-derivative indexes to prevent such entries as AND RIGHT, LEFT see LEFT AND RIGHT. The output of a program segment to produce these inverted references may then be human-edited to eliminate useless or nonsensical references—CENTERS, DAY CARE see DAY CARE CENTERS is useful, but the value of CARE CENTERS, DAY see DAY CARE CENTERS is, to say the least, debatable.

It will be noticed that these see references from inverted forms of multiple-word headings in effect replace the separate permuted lists of used headings included in some thesauri. These references could make up a separate list, of course, but would seem more usefully embodied in the main alphabetical list, where a high percentage of them already appear. Because of the peculiar typography of a permuted KWIC listing, the see references, though more than twice as long, are not only easier to use but probably occupy not much more actual space.

A number of other manipulations may be used to augment the information access provided by the main list by techniques which we have developed (with Edwin B. Brownrigg) and tested on a developmental thesaurus for a newspaper morgue. These include, for example, a listing under each subdivision of those headings to which it has been applied, and cross-references from inverted forms of the subdivisions like those from inverted forms of used headings. It is possible that these extra
aids would be included only in a COM (computer output microform) version of the children's media thesaurus, and not in the printed version, or they could make up a separate list. This procedure would provide a very useful geographic access when combined with the recommendation in regard to indication of indirect subdivision.

A further nicety that might be informative, especially as computer searching of children's materials becomes more widespread, would be to include the frequency of use of used headings up to the time of compilation.

We are now ready to consider production of additional listings other than the main thesaurus listing and the listing of subject subdivisions. A listing consisting simply of used terms is easily produced from the main list, and constitutes basically a handy aid for the experienced cataloger, though it may also serve as a brief hand list for searchers.

A classed list of the headings by Dewey number (remember that more than one class may be assigned each heading) can readily be program generated, in turn, from the list of used headings. It can provide a very useful way to locate headings related to a topic of interest, and the list of used headings of the main alphabetical listing provide a useful index to it.

A second kind of classed listing can be generated from the hierarchical structure implicit in the syndetic (cross-referencing) structure of the main list. Such a list can give the hierarchical structure downward from every term which references a narrower term, or it might be limited to the hierarchy downward only from what W. T. Brandhorst has called head terms—that is, terms which refer to narrower terms but have no references to broader terms.15

If the latter system is adopted, we would suggest a hierarchical numbering for each term in the hierarchical list, somewhat like the numbering used for sections and paragraphs of some manuals and texts. These numbers could then be printed with the associated term in the main list, enabling the main list to serve as an index to the appearances of any term in any position in the hierarchical lists.

**Computer Programs**

We have, and have used in thesaurus production, a complete program set for the production of all of the entry forms and lists proposed here, and have recently rewritten these programs to run on microcomputers.

**Conclusions**

If it were necessary for economic reasons, the first stage of adoption of this proposal could be limited to the production of a new subject-heading list conforming with the American Library Association recommendation for the use of Library of Congress annotated-card practice for the cataloging of children's media. Even the production of the first edition should be less expensive than producing a new edition of Sears and could contain desirable new features. Subsequent editions could be cheaply—and consequently more frequently—produced.
The additional lists and entry forms suggested could be produced as it was convenient to do so, either on COM or in print form, very substantially enhancing information access to children’s media collections without changing the average number or type of headings now used.

It is proposed that the Resources and Technical Services Division Cataloging and Classification Section Cataloging of Children’s Materials Committee consider urging the adoption of this or some similar plan to enhance information access to children’s media and to provide librarians an easy means of reaching conformity with the committee’s earlier recommendations as far as subject cataloging is concerned.

We welcome comments and criticism on the ideas in this paper, as we are currently engaged in the production of a preliminary Thesaurus of Subject Headings for Children’s Materials, based upon Library of Congress subject headings as applied to materials included in The Elementary School Library Collection. Keyboarding is now well under way, and we would hope to produce a preliminary edition in 1983.

REFERENCES


8. These figures were originally provided by Lois Rose, then head, Children’s Literature Section, Subject Cataloging Division, Library of Congress. They were confirmed by telephone on Sept. 28, 1982, by the current head of the section, Jane Marton.


12. Working List.


From: Russell Sweeney, principal lecturer, School of Librarianship, Leeds Polytechnic [Abridged].—The article by Robert H. Hassell in the April/June 1982 issue raises a number of questions about The Proposed Revision of 780 Music. Mr. Hassell attempts a comparison between 19th edition of Dewey Decimal Classification, The Proposed Revision of 780 Music, and what he calls a “performance orientated arrangement.” This arrangement is “the order deemed most advantageous for those seeking to retrieve . . . scores.” In support of this arrangement he makes use of, and produces some tables of statistical comparisons based on 400 scores selected from the British Catalogue of Music (BCM). Since the principles of his arrangement are only “plausible assumptions” and not necessarily universally accepted they need to be examined before considering the comments relating to The Proposed Revision.

The first . . . is that music of interest to a particular performance medium should be kept together. When a score specifies one instrument . . . there is no problem, but when a score specifies more than one instrument a second principle is introduced. The performance orientated arrangement postulates that music for a particular instrument, voice or combination with keyboard should be collocated directly with the same instrument, voice or combination without keyboard, and secondly, music for an ensemble of one particular instrument, with or without keyboard, should be collocated with music for that instrument, groups of scores being arranged in an order of increasing complexity. A third principle is that arrangements should not be interfiled with original music. A fourth principle is that music for wind-string combinations is kept together and collocated with music for wind ensembles. A final principle is that form should not be included in the citation order for music for a single performance media, . . . but that subarrangement should be by composer’s name . . .

Nobody . . . could quarrel with the first principle, but the second principle is debatable. It is possible to argue for the separation of chamber ensembles on the basis of the presence or absence of a keyboard participant, if only on the grounds that the keyboard instrument is not portable and this is a factor in the selection

Editor’s note: Letters sent to the editor for publication in this column cannot be acknowledged, answered individually, or returned to the authors. Whenever space is available in an issue, selected letters will be published, with little or no editing, though abridgment may be required. Letters intended for publication should be typed double-spaced.
of a piece of music by patrons. The case for placing ensembles of one particular instrument with solo music for that instrument is strong and this arrangement is followed by BCM. However, it can also be argued that such music is chamber music and should be grouped with music requiring an ensemble of players. In support of the third principle, it may be that some musicians would scorn the selection of a mere arrangement for his or her instrument, but there would be many who would ... reject this form of musical snobbery. ... There is not likely to be any dissent from the fourth principle, but the final principle is again arguable. A patron requiring a particular Beethoven piano sonata, for example, may be better served if all the sonatas are grouped together and sub-arranged by composer ...

The objections [to the arrangements in The Proposed Revision] seem to be as follows:

1. It does not make any provision for separating collections from original works.
2. It does not make any provision for separating arrangements from original works.
3. By allowing the specification of form in keyboard music it produces an arrangement which is of little performance interest.
4. Separate numbers are provided for harpsichord solos (786.4) and piano solos (786.2).
5. There is no way to separate piano duets for one instrument, from those requiring two instruments, and both are separated from piano solos (786.2).

The responses to these objections are given below

1. Work on The Proposed Revision began in 1974 and the first draft [was] produced in 1975. It was ... revised in 1977 and ... in 1979. During this period the current edition of Dewey Decimal Classification was 18th ed. and Table 1 was applicable. This table contained the subdivision -08 Collections and therefore it was a simple matter to apply this to differentiate between collections and separate works. Unfortunately -08 Collections was removed in 19th ed. published in 1979. It would be a simple matter to restore 780.8 Collections in The Proposed Revision ...

2. The draft ... produced in 1975 made provision for separating arrangements from scores of original instruments, and for specifying the original executant of the arranged piece. However, in response to criticism ... the provision was withdrawn. In any case, it is questionable whether such provision is worth while for the purposes of shelf arrangement. The fact that a particular piece is an arrangement would ... be indicated in the catalogue, and may be indicated on the score.

3. The principles used by Dewey in the Decimal classification, that libraries who do not wish to make use of the detail need not do so, are followed in The Proposed Revision. Whilst it is possible to specify form in the citation order for scores, libraries which prefer a sub-arrangement by composer can easily achieve this by ignoring the form fact.

4. ... the logical outcome would be to deny the existence of subdivisions for clavichords (786.3) and harpsichords (786.4). It is surely up to individual libraries to decide whether or not to place all pieces which can be performed on either harpsichord or piano at one place, presumably with piano music ...

5. The original draft made provision within 786 for duets for two performers on one keyboard instrument. However, with subsequent development of the schedules this provision disrupted the general provisions for synthesis throughout 784-788, and the provision was withdrawn. The question of duets involved the consideration of what constituted chamber (ensemble) music, and for all other instruments duets were classed in 785. Keyboard instruments are a spe-
cial case in that it is possible for two players to perform on the same instrument. With the possible exception of some percussion instruments this is not possible with any other instrument. It was concluded that for The Proposed Revision there should be consistent treatment of duets for any instrument and they should be classed as ensemble music in 785.6-785.9.

Concluding his article, Mr. Hassell makes a number of assertions which are open to challenge. Amongst these is the statement that “the statistics resulting from this study offer no evidence that from a performance standpoint the new scheme provides a better arrangement than that provided by DDC.” The statistics given only have meaning if one accepts the principles of the “performance orientated arrangement.” As this letter attempts to show, these principles are not necessarily acceptable, and the study does not, therefore, provide evidence to justify this assertion, which is . . . based on only a portion of the material which would be classed in 780. . . .

From: Neil Patton, chief, Taxpayer Information and Education Branch, Internal Revenue Service. — This year the IRS is expanding its services to libraries with audio visual materials to help taxpayers prepare their tax returns.

The attached news release describes these materials and our program for libraries. We would appreciate your help in publicizing the materials. If you have any questions, please give us a call. (Contact Johnell Hunter (202-566-6445)

Excerpt from news release: The Internal Revenue Service is offering libraries free audio cassettes, films, and video cassettes of step-by-step instructions on how to prepare tax forms 1040EZ, 1040A, 1040, and Schedules A and B. The 1040EZ form is new this year and is what the name implies: easy. Some 22 million single taxpayers with no dependents, under 65, with income under $50,000 from wages and less than $400 in interest income may be able to use it. All materials feature a series of money-saving tax tips—a suggestion many librarians made in their evaluations of last year’s audio cassettes.

For further information the editor suggests that you contact your local public library and that librarians wishing to offer these services call the number given above.

The editor has been asked to call to your attention two grants that might be of interest to LRTS readers. The new Whitney Fund grant will be awarded to individuals to cover costs associated with the preparation of “bibliographical aids for research.” The Carnegie Fund grant also is to help defray the costs of preparing or publishing bibliographical aids and is available to any official unit of ALA. The deadline for submitting proposals for either grant is the end of February 1983. For further information consult the chair or secretary of the ALA Publishing Committee (312-944-6780).
For the Record

Annual Report of the Decimal Classification
Editorial Policy Committee,
July 1, 1981–June 30, 1982

Margaret E. Cockshutt, Chairperson

The Decimal Classification Editorial Policy Committee (EPC) held its 81st and 82nd meetings at the Library of Congress on October 15–16, 1981 and April 15–16, 1982. The following members were present: Lizbeth Bishop, Barbara Branson, Lois M. Chan, Margaret E. Cockshutt, Betty M. E. Croft, Joseph H. Howard, John A. Humphry, Donald J. Lehnus, Russell Sweeney, and Arnold S. Wajenberg. Others attending both meetings were: John P. Comaromi (Editor of the Dewey Decimal Classification), Lucia J. Rather (Director for Cataloging, Library of Congress), and Judith K. Greene (Secretary). Margaret J. Warren (Assistant Editor) and Walter Curley (Chairperson of the Forest Press Committee) attended the October meetings. Staff members of the Decimal Classification Division (DCD), Library of Congress were present, by invitation, as observers at both meetings.

On October 16 there was also a tripartite meeting of members of the EPC and of the Forest Press Committee (FPC) with the Editor of the Dewey Decimal Classification (DDC) and his staff. In addition to those named above, this tripartite meeting was attended by Harold Bonner, Richard Gardner, Michael O’Brien, and Thomas Sullivan, of the Forest Press Committee; William J. Welsh (Deputy Librarian of Congress); and Mildred Dugan (Forest Press).

The EPC acted on the following matters:

1. It accepted 301–307 Sociology, Expanded Version Based on Edition 19 of the DDC, together with its Index, Equivalence Tables to link Edition 18 back and forth with the expanded version of Edition 19, and the “Manual of Application for the 301–307 Expansion.” The expanded schedule was produced in response to requests from librarians and library organizations, and to meet very recent needs in literary warrant. Copies are being distributed free of charge to purchasers of Edition 19, and to users of the Abridged Edition who request it, as announced in special flyers and in advertisements in library journals.

2. The EPC recommended to the Forest Press Committee that the Introduction to the Manual on the Use of the Dewey Decimal Classification be published, with some editorial refinement. This recommendation constituted the EPC’s final approval of the Manual, with publication expected in the Summer 1982.

3. The Committee reviewed several proposals and drafts of editorial rules for Edition 20. The DCD and the EPC hope that revised editorial rules will make notes and other features of the schedules of Edition 20 clearer and
more informative for DDC users. Because of the importance of the rules for ultimate ease of use of the DDC by classifiers, the EPC asked that some proposals be studied further by the Division.

4. A resolution was passed in thanks to Margaret J. Warren who retired early in 1982 as Assistant Editor of the Dewey Decimal Classification. The Committee recognized with gratitude her work on the development of Editions 18 and 19, and her contribution to the discussions and planning for Edition 20. The EPC noted her “careful and deliberate analysis and logical arrangement of diverse fields of knowledge” and her “quite remarkable and splendid creative work” in her editorial work on the DDC.

In the October 1981 tripartite meeting, members of the EPC and FPC joined with the Editor and staff members of the Decimal Classification Division in a wide-ranging discussion of the future of the DDC. There was no attempt to reach decisions at this meeting. Rather, it was intended to be a forthright exploration of the basic issues to be faced.

Inevitably, the most thorny problem is that of ‘keeping pace with knowledge’ versus ‘integrity of numbers’. It was recognized once again that some specific schedules must be revised completely (‘phoenix’ schedules) when a poor intellectual structure causes classifiers many practical problems in applying a schedule, or when radical changes in a subject over time make an existing schedule obsolete. Such changes may also be necessary in response to a changed societal perception of a subject as, for example, divorce which used to be seen as a subject in religion or ethics but is now seen as a social problem. At the same time it was recognized that phoenix revisions, which re-use numbers without regard to their former meaning, cause extensive problems for classifiers who must fully implement the new schedule and thus reclassify old material to maintain collocation; or reject the phoenix schedule and thus perpetuate an increasingly obsolete and difficult arrangement; or maintain the old arrangement and implement the phoenix for new material only, and thus negate the collocation to the detriment of browsers.

To avoid the cataclysmic impact of many substantial changes in a new edition, continuous revision was considered. This could be in the form of interim separate publications, such as the new 301–307 Sociology, Expanded Version, and/or publication of changes in Decimal Classification Additions, Notes and Decisions (DCN). If there were immediate implementation of such changes by national bibliographic agencies, libraries might have difficulty in determining the meaning of new or changed notations and in identifying which versions of an edition were being applied centrally. However, without the early centralized implementation of necessary revisions between editions, the impact on libraries would be much heavier at the time of a new edition. Also, because DDC is a unitary system which covers the whole field of knowledge, with the various classes interconnected through the “add” mechanism, the fragmented publication of many revisions could destroy the unity of the whole and make the application of DDC inefficient and difficult for classifiers. Clearly, regular and frequent accumulations of schedules and/or indexes would be necessary in a program of continuous revision. The possibility of a loose-leaf format was also discussed.

If continuous revision with interim publication and application is adopted, the form of payment for the DDC is another critical question for both libraries and Forest Press, the publisher. Should there be an annual subscription payment, which might cover both DCN and interim separates? Would such a subscription also include any accumulations? Should a subscription to a DDC online service also be possible? While no final decisions on revisions, formats and forms of payment were made at the tripartite meeting, agreement on general directions towards change seemed to be emerging.
At both the October and April meetings, the EPC had extensive discussion without action on other matters:

1. The Committee discussed the present segmentation policy applied by national bibliographic agencies, in terms of the relationship between the full number for a concept in the Abridged Edition and the number indicated by the first prime mark in the segmentation rules applied to Edition 19 numbers. At present the first prime mark indicates only a column number in the Abridged Edition, i.e., the possible standard subdivisions, area numbers and other tables are not applied. While it may seem desirable for the first prime mark to indicate the full Abridged number, this change would require some technical revisions to the 11th Abridged Edition. Since DC&F is not currently distributed to Abridged Edition users as a vehicle for the dissemination of interim revisions, a decision on a revised policy for segmentation was deferred.

2. The EPC discussed at length the current index to DDC Edition 19 and ways in which it could be enriched and improved. Selected descriptors for the Library of Congress Subject Headings (LCSH) might be included in the index, as perhaps being closer to popular language, for the assistance of classifiers. The Committee also considered a preliminary paper by Gregory R. New (DCD staff) proposing a two-level index: a shortened index to a few of the most important or difficult numbers for specific subjects; and a full relative index to a selected list of key generic terms. Much more study is necessary before there is a policy decision on changes in the index.

3. The Committee examined the first draft of a possible phoenix schedule of 370 Education. This was prepared by Margaret J. Warren at the EPC’s request, in order to take full advantage of Warren’s expertise before her retirement. The structure of the schedule is based in part on the Bliss Bibliographic Classification, 2nd edition. While no firm decision has been made on any phoenix schedules for Edition 20, 370 has been discussed as a candidate for phoenix revision. An examination of this draft showed once again the close relationship between the 370s and standard subdivision 07. Since a change in any of the standard subdivisions would have major ripple effects through the entire classification, it was agreed that the content and structure of Table 1 must be considered per se, and that this study must be done before a decision can be made about the nature and extent of the future revision of 370.

4. The EPC considered various recent articles, reviews and notes pertaining to the DDC and their implications for future work by the Committee. It also noted the Reading List: Dewey 19 prepared in March 1982 by the (British) Library Association Library.

In Committee business, Margaret E. Cockshutt was re-elected as Chairperson to October 1983. The next EPC meeting was scheduled for November 4-5, 1982.
INSTRUCTIONS TO AUTHORS

Editor's note: Because of the recent appearance of the thirteenth edition, revised and expanded, of The Chicago Manual of Style, the "Instructions to Authors" are now being revised. The revised instructions will appear in the April/June issue of this volume.
Everybody's talking about on-line catalogs

What about on-time?

GRCCOM

A name that stands for on-time!

GENERAL RESEARCH CORPORATION
A SUBSIDIARY OF FLOW GENERAL, INC.
P.O. Box 6770, Santa Barbara, CA 93111-0770
Phone (805) 964-7724
Contact: Vera Fessler

MUTIPRINT YOUR OWN BOOK CATALOG
Union List, Directory or Bibliography in fifteen working days. From computer photocomposition, line printer or other line copy ready for camera. Write with your specifications including number of copies, number of pages and page size for Quotation, free Production Guide and Sample Book Catalog. Mention LRTS for favorable price. Multiprint, Inc., 28 West 23rd Street, New York, NY 10010, (212) 924-1100.
CLASS LIBRARY SOFTWARE

The Affordable Alternative

CHECKMATE
—Serials Control—
- Searching
- Accounting
- Title and Subject Lists
- Subscription Expiration Alert

GOLDEN RETRIEVER
—File Management—
- Current Awareness
- Catalogs
- Bibliographies
- Searching
- Sorting
- Indexes
- Searching

Both systems require a microcomputer with at least 64K of CPU memory, and either CP/M, PC DOS, or TRS-DOS. The program is written in Microsoft BASIC. Demonstration disks of either system are available for $21.50. A Users Manual is also sold separately for $31.50.

For more information, please contact one of the following distributors:

CLASS 1415 Koll Circle, Suite 101, San Jose, CA 95112-4689 (408) 289-1756

Capital Systems Group, Inc. (CSG) 11301 Rockville Pike, Kensington, MD 20895 (301) 881-9400

Gaylord Brothers, P.O. Box 4901, Syracuse, NY 13221 (315) 457-5070

"1/2 Million Ways To Creative Teaching"

Total Information Service on Non-Print Media Indexes provides access to over one-half million annotated media titles including 16mm films, 35mm filmstrips, videotapes, slides, records, audio tapes and others. Also includes multimedia indexes in Special Education, Psychology, Health & Safety, Vocational & Technical—many others.

nicem National Information Center for Educational Media
University of Southern California Research Annex — Suite 301 3716 South Hope Street Los Angeles, California 90007
CALL TODAY AND PLACE YOUR ORDER:
Phone: 800-421-8711 in California (213) 743-6681
"The profession is fortunate to have this periodical."

**SERIALS REVIEW**

"Serials Review can improve the operation of most public and academic library serial departments...a much improved periodical and one that should be treasured by serials librarians as a useful guide to both new and older titles in the field....The profession is fortunate to have this periodical. Recommended."

—Bill Katz


"The broad scope of this journal makes it a reference and collection development tool for all but the smallest libraries."

—Arba '81


Send for a free sample copy of SR. While you're at it, request a sample of SR's sister publication, Reference Services Review.

pierian press  p.o. box 1808  ann arbor, mi  48106  usa

---

**It's Time To Try A New ROUTE.**

Introducing ROUTE, The Newest Module In Faxon's LINX System.

Now there's a computerized system for creating, updating and printing journal routing slips. The ROUTE system.

With ROUTE, all the time-consuming details of managing routing lists are taken care of for you. ROUTE keeps track of who gets what publications, it lists readers by name, location, address and other identifiers you specify. It can prioritize routing order by title or other parameters, and it produces all your routing slips for you.

☐ Show Me The ROUTE.

I'm looking for a way to manage periodical distribution more efficiently. Send me all the facts on your ROUTE system right away.

MAIL

ADDRESS

CITY

STATE

ZIP

ROUTE also gives you valuable management and statistical reports. It simplifies the procedure for changing reader records, and much more.

For the full story on easy-to-use, easy-to-learn ROUTE, part of Faxon's LINX serials management system, just return the coupon or call 617-329-3350, outside of Boston, call 800-225-6055.

faxon

The Art and Science of Serials Management.
SERVING PEOPLE: THE PUBLIC LIBRARY
TODAY AND TOMORROW

Plan to attend the Public Library Association’s National Conference
Baltimore, Maryland -- March 23-26, 1983

Two hundred hours of planned activity - 15 hours of exhibits, 70 hours of theme tours,
109.5 hours of programming, & 5.5 hours of socials - will interest
the public librarian, trustee, educator, library school student, & friend of the public library.

Take this opportunity:

- to hear the notable & the experts
- to learn new skills
- to broaden professional knowledge
- to use Placement Services
- to increase national visibility of the public library
- to meet old friends & make new ones
- to experience a delightful city

For more information write or call:
Shirley Mills-Fischer, Executive Director, PLA/ALA, 50 E. Huron St., Chicago, IL 60611
312/944-6780

A breakthrough in label technology has created a Cold Application
pressure sensitive label with a foil backing that retains the shape of the
application. These labels are smudgeproof and flexible, with a perma-
nent adhesive that really sticks to books. Perma/Seal II spine labels
and OCLC SL6D and SL4D labelsets are available.

Send for your free
Foil Back Label sample
and a copy of our new
Library Supplies
Catalog.

NAME: ___________________________
ADDRESS: ______________________
CITY: __________________________
STATE: __________ ZIP: __________

☐ SAMPLE ☐ CATALOG
DOES YOUR SUBSCRIPTION AGENCY OFFER YOU... 
MORE?

MORE... than 160,000 foreign and domestic serials, including continuations and annuals?

MORE... personalized service and flexibility from regional offices?

MORE... easily handled claiming, with the least time-consuming claim system:
A. A simple list
B. A monthly Claim Checker
C. A Missing Copy Bank?

MORE... innovation, with the world's first international online data communications system for serials?

MORE... professionals with the knowledge and concern to make your serials program work better?

We provide the most...MORE often.
EBSCO SUBSCRIPTION SERVICES
The serials professionals.

P. O. Box 1943
Birmingham, Alabama 35201-1943
(205) 252-1212
Telex: 5-9717
We've collected some of the great issues of our times on archival quality 35mm microfilm. The news and opinion as presented by preeminent publishers worldwide, for perusal and scholarly research.

The Times of London, and its supplements, available from 1785, or as the supplements were established by Times Newspapers of Great Britain. These newspapers share a tradition of editorial integrity and authority.

Le Monde archives dating from 1944, including its companion publications, Le Monde De L’Education and Le Monde Diplomatique, are considered to be among the most intellectual of newspapers.

The Washington Post, available from 1878, chronicles the United States’ emergence to global prominence as reflected by domestic and international news.

Research Publications offers 22 other important world newspapers in microfilm collections, including the recent addition of Neue Zürcher Zeitung and the South China Morning Post. These collections, and those of our periodicals and serials, are produced to uncompromising standards to become a permanent resource to scholarly research.

We’ve got the issues where the answers are found, and the indexes to provide ready access. And we offer a number of options for the acquisition of these and other microfilm collections.

If you want to know more, try Research Publications, Inc.

Send me more information on your newspaper collections, in particular:

☐ The Times of London
☐ Le Monde
☐ The Washington Post
☐ The catalog on your other newspapers.
☐ Keep me informed about new offerings.

Name ____________________________
Title ____________________________
Street ___________________________
City __________________ State ______ Zip ______ Phone ______
Good Directions from ALA

**Cartographic Materials**
A Manual of Interpretation for AACR2
Prepared by the Anglo-American Cataloguing Committee for Cartographic Materials
Hugo L. P. Stibbe, General Editor
Vivien Cartmell and Velma Parker, Editors

This manual contains everything necessary to give bibliographic descriptions of maps, globes, atlases, and other cartographic materials in accordance with AACR2. It indicates preferred applications of AACR2 together with the practices of the national libraries and archives of Australia, Canada, New Zealand, Great Britain, and the United States.

x, 258 pages, illustrated 0363-0
$40.00 Cloth

**The Bibliographic Record and Information Technology**
Ronald Hagler and Peter Simmons

This innovative text aims to introduce novices to computer functions in the context of bibliography, to deal with manual and automated bibliographic databases as part of a single coordinated system of bibliographic control, and to analyze the technical, administrative, and economic aspects of a rapid transition from older rules and practices to newer ones.

xx, 346 pages 0376-2 $25.00 Cloth

**Cataloging Machine-Readable Data Files**
An Interpretive Manual
Sue A. Dodd
Foreword by Judith S. Rowe

The only work of its kind, this manual provides the guidance which data producers, data archivists, and data librarians need to supply consistent bibliographic information for the machine-readable data files they service. Dodd explicates the rules for cataloging MRDF from the ninth chapter of AACR2, offering examples and revealing the "how and why" of cataloging. Included is an extremely valuable outline of the steps for cataloging the microcomputer programs that are now found in many public and school libraries.

xx, 248 pages, illustrated 0365-7
$35.00 Paper

**AACR2 Decisions and Rule Interpretations**
Second Edition
Compiled by C. Donald Cook with the assistance of Glenna E. Stevens

Published by the Canadian Library Association; distributed in the U.S. by ALA. This consolidation appears on loose-leaf pages so users may have the option of interleaving them in a copy of AACR2 or binding them as a separate volume.

3281-9 $50.00

**American Library Association**
50 East Huron Street, Chicago, Illinois 60611
At Midwest Library Service, We Take The Team Approach To Assist Your Library

Carl Dorr  Pam Rodgers  Forrest E. Link  Carol Lehmkuhl
(SOUTHEAST)

Kevin P. Coyle  Linda Market  Kim S. Anderson  Glenda Ward
(MIDWEST)

Lawrence D. Nagel  Pat Hamil
(WEST)

To best serve your needs, we have formed five problem-solving service teams to help take the hassle out of book-buying. Each team is composed of a Sales Representative in the field and a Customer Service Representative in our home office.

Once alerted by your phone call made on our Toll-Free WATS Line, 1-800-325-8833, (Missouri customers, please call COLLECT 0-314-739-3100) your problem-solving team, geographically assigned to your library, goes into action immediately.

It is another facet of Midwest Library Service's tradition of excellence.

May we have the privilege of serving your library?

Midwest Library Service
11443 St. Charles Rock Road
Bridgeton, Mo. 63044

"23 Years of Service To College and University Libraries"
Imagine. The LC National Union Catalog from 1953 to 1982 and the Catalog of Books Represented by LC Printed Cards from 1898 to 1952 in convenient, easy to use microfiche form.

The equivalent of 693 volumes at a single convenient microfiche reader station, with sharp, clear entries enlarged twice the size of printed editions, outlasting printed editions, plus saving shelf space.

Now, exclusively and for the first time ALS offers this indispensable reference source at a 75% savings over printed alternatives.

This immense collection, known as the standard for cataloging, offers full bibliographic descriptions. It will always be invaluable to any library for cataloging, acquisitions, reference work and many other purposes, and indispensable where research is done for author bibliography, verifying titles, bibliographic information, historical notes, and location of copies. And for $495 the entire (1953-1982) set of LC Audiovisual and Music Catalog Supplements to the NUC may be added.

Also, the complete LC Subject Catalog (1950-1982) is now available on microfiche for only $3,475. This comprehensive collection provides subject access to most of the world’s output of significant books and serials. Current subscriptions, individual quinquenniums and annuals available for all LC Catalogs.

Prices slightly higher for non-U.S. libraries.

Use coupon or call 617-470-0610 for complete details.

<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
<th>L/R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Organizations

Address

City/State/Zip Code

Please rush me complete information on LC micropublications:

☐ National Union Catalog
☐ Audiovisual and Music Catalogs
☐ Catalog of Books
☐ Subject Catalog
☐ Please have your representative phone me at ( )

Advanced Library Systems Inc.
93 Main Street
Andover, MA 01810, U.S.A.
(617) 470-0610
BNA EXPANDED AUTHORITY CONTROL SERVICES.
HELPING YOU REALIZE THE SAVINGS 
YOUR MARC CONVERSION PROMISED.

Converting your library’s catalog to machine-readable records was an immense task. Now, converting those records to AACR2 cataloging rules creates a new challenge. Split catalogs, "it’s-as-it’s" techniques and extensive cross reference structures are costly to maintain and difficult to use. Certainly automated authority control is the answer.

We at Blackwell Technical Services are specialists in automated authority control. You simply provide your machine readable database. We run your records against the continually updated AACR name and subject authority files.

The advantages are obvious. There is a single, consistent authority for your entire catalog. You may choose comprehensive authority control or one of the following:

Subject Authority Control. This well known Blackwell service updates subject headings to latest AACR practice. We have updated our authority file from the printed lists, completing it through the ninth edition plus the 1979 supplement.

Name Authority Control. Because 75% of AACR changes concern names, we now offer this sophisticated service to painlessly convert name headings.

Our authority control services update headings, correct spelling and capitalization errors, correct AACR tags and delimiters, create cross reference and eliminate blind cross references. The few exceptions which cannot be corrected automatically are reviewed in context by experienced Blackwell authority control editors who apply appropriate changes. Then we produce your catalogs on microfilm, microfiche, or paper or we deliver a tape copy of your revised files.
Acronyms, Initialisms, and Abbreviations Dictionary
8th ed. Vol. 1, Acronyms, Initialisms, and Abbreviations Dictionary. 254,000 entries. 1,701 pp. in 2 vols. 1982. $98.00/set. (SO) Vol. 2, New Acronyms, Initialisms, and Abbreviations. Two inter- edition supplements. Paper. 1983-84. $92.00/sub. (SO) Vol. 3, Reverse Acronyms, Initialisms, and Abbreviations Dictionary. Contains entries in Volume 1 but is arranged by word or term, with acronym as definition. 1,700 pp. in two volumes. $120.00/set. (SO)

Classified Directory of Artist's Signatures, Symbols, and Monograms
2nd ed. 7,500 entries, each providing one or more facsimile signatures, plus concise biographical data and exhibition information. 873 pp. 1982. $195.00. (SO)

First Names First

Linguistic Atlas of the Upper Midwest
The study is an exploration of regional patterns in American language in Minnesota, Iowa, North Dakota, South Dakota, and Nebraska. 829 pp. in 3 vols. Published by the University of Minnesota Press, 1973-76. Available from Gale. $96.00/set.

LC Classification Sched
Indexes to P Schedules and Additions and Changes 1978. Gale has now indexed the P subclasses, whi language and literature. $500.00/set. (Indexes are separately.) (SO)

Modifiers
A Unique, Compendious C of More Than 16,000 Engli tives Relating to More Th Common and Technical Nouns. 1st ed. Modifiers for those adjectives in English not simply formed by adding to the noun form. This work solves the problem faced by the writer who needs to find unusual adjectival forms. 203 pp. 1982. $48.00. (SO)

Pseudonyms and Nicknames Dictionary
2nd ed. A convenient, one-volume compendium of information on over 50,000 pseudonyms, nicknames, and other assumed names used by over 40,000 actual personalities. Includes historical individuals. 995 pp. 1982. $135.00. (SO)

Word for Word
A Dictionary of Etymological Cognates. 1st ed. Explains in everyday English the differences and similarities between cognates—words in modern English that look and sound quite different from one another yet share a common ancestor. 454 pp. Published by Verbatim, 1982. Distributed to libraries by Gale. $42.00. (SO)

(SO) These titles are available at Gale’s 5% Standing Order discount.
All Gale books are sent on 90-day approval. Deduct 5% if you send check with order.
Customers outside the U.S. and Canada add 10% to prices shown.

Gale Research Co. Book Tower • Detroit, MI 48226